

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

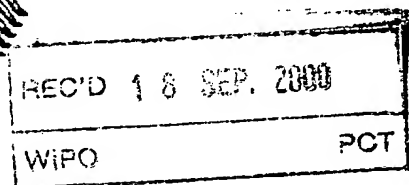
**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**



**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)



16800 U.S. PTO  
10/069353



**Prioritätsbescheinigung über die Einreichung  
einer Patentanmeldung**

EP 00/08013

**Aktenzeichen:**

199 57 268.2

**Anmeldetag:**

29. November 1999

**Anmelder/Inhaber:**

Bayer Aktiengesellschaft,  
Leverkusen/DE

**Bezeichnung:**

Nucleinsäuren, die für Enzymaktivitäten der  
Spinosyn-Biosynthese codieren

**Priorität:**

27.08.1999 DE 199 40 596.4

**IPC:**

C 12 N, C 12 P

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Anmeldung.

München, den 08. Juni 2000  
Deutsches Patent- und Markenamt  
Der Präsident  
Im Auftrag

Wehner

Aufbauend auf Untersuchungen zum Einbau von  $C^{13}$ -markiertem Butyrat oder Isobutytrat konnte gezeigt werden, dass die Biosynthese eines Polyketids einem Polyketid-Biosyntheseweg folgt (Nakatsukasa et al., 1980). Durch multifunktionelle Enzyme, den sog. Polyketidsynthasen, kettigen Säurebausteinen wie Acetat, Propionat oder Butyrat, die die verwandten Fettsäuresynthasen (FAS's) katalysieren, sind die Kondensationsschritte der als CoA-Thioester aktivierten Bausteine nach jedem Kondensationsschritt eine vollständige Reduktion der wachsenden Polyketidkette, die entstehenden  $\beta$ -Oxoester durch Dehydratation und Enoylreduktion katalysieren, können PKS's diese Schritte auslassen. Modulare Typ I PKS's bestehen aus einem multifunktionalen Protein. Iterative Typ II PKS's stellen dies aus weitgehend monofunktionalen Proteinen dar.

Die enzymatischen Aktivitäten von modularen Typ I PKS's lassen sich in bestimmten Modulen zusammenfassen. Hierbei trägt ein Modul eine oder mehrere enzymkatalytisch aktiven Domänen, die zu einer Verlängerung der Polyketidkette um eine biosynthetische Verlängerungseinheit führen. In einer solchen Domäne handelt es sich um eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne, eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne. Ein Modul kann auch eine Ketoreduktase-, eine Dehydratase- und eine Thioesterase-Domäne tragen. Ein sog. Linearer Beginn der Biosynthese steht kann von den genannten Domänen gebildet werden. Eine enzymatisch inaktive  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne, eine Polyketidsynthase-Domäne, umfasst jeweils eine dieser genannten Aktivitäten.

Aufgrund der potenten insektiziden Wirkung sowie der Bedeutung der Spinosyne besteht ein großes Interesse, die genetischen Informationen der Biosynthese zu entschlüsseln.

Aufbauend auf Untersuchungen zum Einbau von  $C^{13}$ -markiertem Acetat, Propionat, Butyrat oder Isobutytrat konnte gezeigt werden, dass die Biosynthese von A83543 einem Polyketid-Biosyntheseweg folgt (Nakatsukasa et al., 1990). Polyketide werden durch multifunktionelle Enzyme, den sog. Polyketidsynthasen (PKS's) aus kurz-

5 kettigen Säurebausteinen wie Acetat, Propionat oder Butyrat aufgebaut. Ähnlich wie die verwandten Fettsäuresynthasen (FAS's) katalysieren sie decarboxylierende Polykondensationsschritte der als CoA-Thioester aktivierten Bausteine. Während FAS's nach jedem Kondensationsschritt eine vollständige Reduktion der intermediär an der wachsenden Polyketidkette entstehenden  $\beta$ -Oxoester durch Ketoreduktion, Dehy-

10 dratation und Enoylreduktion katalysieren, können PKS's bestimmte Reduktionsschritte auslassen. Modulare Typ I PKS's bestehen aus einem oder mehreren großen multifunktionalen Proteinen. Iterative Typ II PKS's stellen dagegen einen Komplex aus weitgehend monofunktionalen Proteinen dar.

15 Die enzymatischen Aktivitäten von modularen Typ I PKS's lassen sich zu sogenannten Modulen zusammenfassen. Hierbei trägt ein Modul eine Anordnung von drei enzymkatalytisch aktiven Domänen, die zu einer Verlängerung der wachsenden Polyketidkette um eine biosynthetische Verlängerungseinheit führen. Bei diesen Domänen handelt es sich um eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-

20 Domäne, eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne. Ein Modul kann auch eine Ketoreduktase-, eine Dehydratase-, eine Enoylreduktase- und eine Thioesterase-Domäne tragen. Ein sog. Ladungsmodul, das am Beginn der Biosynthese steht kann von den genannten Domänen lediglich eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein-Domäne tragen, sowie

25 eine enzymatisch inaktive  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne. Eine Polyketidsynthase-Domäne umfasst jeweils eine dieser genannten enzymatischen Aktivitäten.

Aufgrund der potenten insektiziden Wirkung sowie der bemerkenswerten Struktur

30 der Spinosyne besteht ein großes Interesse, die genetischen Informationen für deren Biosynthese zu entschlüsseln.

Gegenstand der Erfindung sind Nucleinsäuren, welche zumindest eine Region umfassen, die für eine Enzymaktivität codiert, welche an der Biosynthese von Spinosynen beteiligt ist.

5

Die vorliegende Erfindung stellt ein Cluster von offenen Leserahmen (ORF's) bereit, deren Translationsprodukte an der Biosynthese von Spinosynen beteiligt sind. Weiterhin werden zusätzliche Gene bzw. ORF's bereitgestellt, die außerhalb des ca. 120 kb großen Spinosyn-Biosynthesecusters liegen, und deren Translationsprodukte an der Rhamnose-Zuckerbiosynthese beteiligt sind.

10

Bei den erfindungsgemäßen Nucleinsäuren handelt es sich insbesondere um einzelsträngige oder doppelsträngige Desoxyribonucleinsäuren (DNA) oder Ribonucleinsäuren (RNA). Bevorzugte Ausführungsformen sind Fragmente genomischer DNA und cDNA's.

15

Der Ausdruck "zumindest eine Region", wie er hierin verwendet wird, bedeutet, dass die erfindungsgemäße Nucleinsäure eine oder mehrere Sequenzen umfassen kann, welche jeweils für einzelne Aktivitäten codieren, die Schritte bei der Synthese von Spinosynen durchführen. Es werden demnach auch Nucleinsäuren als erfindungsgemäß betrachtet, die nur für eine einzige Enzymaktivität der Spinosyn-Biosynthese codieren.

20

Der Ausdruck "Enzymaktivität", wie er hierin verwendet wird, bedeutet, dass ausgehend von den hierin betrachteten Nucleinsäuren zumindest derjenige Teil eines vollständigen Enzyms exprimiert werden kann, der noch die Katalyseeigenschaften des Enzyms ausübt.

25

Insbesondere codieren die erfindungsgemäßen Nucleinsäuren für Enzymaktivitäten von Polyketidsynthasen, Methyltransferasen, Epimerasen, Glycosyltransferasen,

30

Aminotransferasen, Dimethyltransferasen, Reduktasen, Dehydratasen und/oder Cyclisierungsenzymen.

5 Bevorzugt handelt es sich bei den erfindungsgemäßen Nucleinsäuren um DNA-Fragmente, die genomischer DNA von *S. spinosa* entsprechen.

Besonders bevorzugt umfassen die erfindungsgemäßen Nucleinsäuren zumindest eine Sequenz ausgewählt aus

- 10 (a) den Sequenzen gemäß SEQ ID NOS: 1, 2, 3, 4, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 52 oder 54,
- 15 (b) zumindest 14 Basenpaare langen Teilsequenzen der unter (a) definierten Sequenzen,
- (c) Sequenzen, welche an die unter (a) definierten Sequenzen hybridisieren,
- 20 (d) Sequenzen, welche eine zumindest 70 %ige, bevorzugt eine 80 %ige, besonders bevorzugt eine 90 %ige Identität zu den unter (a) definierten Sequenzen aufweisen,
- 25 (e) Sequenzen, welche zu den unter (a) definierten Sequenzen komplementär sind, und
- (f) Sequenzen, welche aufgrund der Degeneriertheit des genetischen Codes für dieselbe Aminosäuresequenz kodieren wie die unter (a) bis (d) definierten Sequenzen.
- 30

Der Ausdruck "hybridisieren", wie er hierin verwendet wird, beschreibt den Vorgang, bei welchem ein einzelsträngiges Nucleinsäuremolekül mit einem komplementären Strang eine Basenpaarung eingeht. Auf diese Weise können beispielsweise ausgehend von genomischer DNA aus Organismen, die phylogenetisch mit *S. spinosa* verwandt sind und die Fähigkeit der Biosynthese von Spinosynen besitzen, DNA-Fragmente isoliert werden, welche dieselben Eigenschaften wie die aus *S. spinosa* isolierten Fragmente aufweisen.

Bevorzugte Hybridisierungsbedingungen sind nachstehend angegeben: Hybridisierungslösung: 5 x SSC; Blocking Reagents (Roche Diagnostics GmbH, Mannheim, Deutschland), 1 %; N-Lauroylsarcosin, 0,1 %; SDS (Sodiumdodecylsulfate) 0,02 %; Hybridisierungstemperatur: 60°C; erster Waschschrift: 2 x SSC bei 60°C; zweiter Waschschrift: 2 x SSC bei 60°C; bevorzugt zweiter Waschschrift: 0,5 x SSC bei 60°C; besonders bevorzugt zweiter Waschschrift: 0,2 x SSC bei 60°C.

Der Grad der Identität der Nucleinsäuren wird vorzugsweise bestimmt mit Hilfe des Programms GAP aus dem Programmpaket GCG (Devereux et al., 1984), Version 9.1 unter Standardeinstellungen.

Besonders hervorgehoben werden Nucleinsäuren, die

- (1) entweder alle Sequenzen, die für Schritte der Forosamin- und Trimethyl-Rhamnose-Biosynthese codieren, umfassen, insbesondere die Sequenzen gemäß SEQ ID NOS: 4 und 51, oder
- (2) alle Sequenzen, die für Schritte der Polyketidsynthese codieren, umfassen, insbesondere die Sequenzen gemäß SEQ ID NOS: 5 und 6, oder
- (3) alle Sequenzen, die für alle Schritte der Forosamin-, Trimethyl-Rhamnose- und Polyketidsynthese codieren, umfassen, insbesondere die Sequenzen gemäß SEQ ID NOS: 1, 2, 3 und 51.

Alle zur Spinosyn-Biosynthese oder zur Synthese von Vorstufen, wie sie nachstehend definiert sind, benötigten DNA-Sequenzen können sich somit auf einem einzelnen Vektor befinden. Diese Nucleinsäuren können aber auch auf zwei oder mehreren Vektoren vorliegen und gleichzeitig oder nacheinander in einer Wirtszelle exprimiert werden.

Alle ORF's der erfindungsgemäßen Nucleinsäuren können von ihren eigenen Promotoren oder von heterologen Promotoren angeschaltet werden.

Gegenstand der vorliegenden Erfindung sind auch die regulatorischen Regionen, welche natürlicherweise, d.h. im Ursprungsorganismus *S. spinosa*, die Transkription der erfindungsgemäßen Nucleinsäuren kontrollieren.

Der Ausdruck "regulatorische Regionen", wie er hierin verwendet wird, bezieht sich auf Promotoren, Repressor- oder Aktivator-Bindungsstellen, Repressor- oder Aktivatorsequenzen, und Terminatoren. Ferner sind genetisch mobile Elemente, welche natürlicherweise, d.h. im Ursprungsorganismus *S. spinosa* vorkommen, ebenfalls von diesem Ausdruck umfasst. Solche genetisch mobilen Elemente können transposable oder mobilisierbare Elemente oder funktionelle Teile davon, IS-Elemente oder andere Insertionselemente sein. Weiterhin sind auch amplifizierbare DNA-Elemente (Amplifiable Units of DNA, AUD; Fishman and Hershberger, 1983), welche natürlicherweise, d.h. im Ursprungsorganismus *S. spinosa* vorkommen, von diesem Ausdruck umfasst. Die Erfindung betrifft auch jede Kombination dieser regulatorischen Regionen untereinander oder mit heterologen DNA-Fragmenten, wie z.B. Promotoren, Repressor oder Aktivator-Bindungsstellen, transposablen, mobilisierbaren oder transduzierbaren Elementen.

Gegenstand der vorliegenden Erfindung sind weiterhin DNA-Konstrukte, die zumindest eine erfindungsgemäße Nucleinsäure und einen heterologen Promotor umfassen.

Der Ausdruck "heterologer Promotor", wie er hierin verwendet wird, bezieht sich auf einen Promotor, der im Ursprungsorganismus nicht die Expression des betreffenden Gens (ORF's) kontrolliert.

5 Die Auswahl von heterologen Promotoren ist davon abhängig, ob zur Expression pro- oder eukaryotische Zellen oder zellfreie Systeme verwendet werden. Ein bevorzugtes Beispiele für einen heterologen Promotor ist der Promotor des *mel*-Gens aus dem Vektor pIJ702 (The John Innes Foundation, Norwich, UK 1985). Die heterologe Expression kann z.B. eingesetzt werden, um zu einer Steigerung der Produktion von Spinosyn im Vergleich zum natürlichen Spinosyn-Produzenten zu gelangen.

10 Gegenstand der Erfindung sind ferner Vektoren, die zumindest eine der erfindungsgemäßen Nucleinsäuren enthalten. Als Vektoren können alle in molekularbiologischen Laboratorien verwendeten Phagen, Plasmide, Phagmide, Phasmide, 15 Cosmide, YACs, BACs, PACs, künstliche Chromosomen oder Partikel, die für einen Partikelbeschuss geeignet sind, verwendet werden.

Bevorzugt sind BAC-Vektoren. BAC-Vektoren (Bacterial Artificial Chromosome) sind entwickelt worden zur Klonierung von großen DNA-Fragmenten (Shizuya et al., 20 1992). Es handelt sich um "single-copy" Plasmide mit einem F-Faktor Origin, die DNA-Fragmente mit einer durchschnittlichen Größe von 120 Kilobasenpaaren (kb) tragen können. Sie sind replizierbar in *Escherichia coli*. Der BAC-Vektor pBeloBAC11 (Kim et al., 1996) trägt einen T7 und einen SP6 Promotor, welche die Klonierungsstelle flankieren und als Startbereich für Sequenzierungsprimer sowie 25 zur Generierung von RNA-Transkripten verwendet werden können.

Besonders bevorzugt sind die am 20. August 1999 bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D-38124 Braunschweig, in Übereinstimmung mit den Bestimmungen des Budapester Ver- 30 trages unter den Hinterlegungsnummern DSM 13010, DSM 13011 und DSM 13012 hinterlegten BAC-Shuttleklone, die Gegenstand der vorliegenden Erfindung sind.



Die hinterlegten BAC-Shuttleklone P11/G6, P8/G11 und P11/B10 tragen jeweils ein mindestens 100 kb großes DNA-Fragment aus *S. spinosa*. Die Klone P11/G6 und P11/B10 tragen jeweils einen Teil der Nucleinsäuresequenz gemäß SEQ ID NO: 4, sowie die angrenzenden vollständigen Nucleinsäuresequenzen gemäß SEQ ID NOS: 5 und 6, sowie einen an die Nucleotidsequenz gemäß SEQ ID NO: 6 3'-angrenzende DNA-Bereich (Abb. 7). Der Klon P8/G11 trägt einen Teil der Nucleinsäuresequenz gemäß SEQ ID NO: 6, die vollständigen Nucleinsäuresequenzen gemäß SEQ ID NOS: 5 und 4, sowie einen an die Sequenz gemäß SEQ ID NO: 4 3'-angrenzenden DNA-Bereich (Abb. 7).

In gleicher Weise sind auch PAC- und alle anderen funktionell gleichwertigen Vektoren, die es erlauben, große DNA-Fragmente, insbesondere solche DNA-Fragmente, die größer als 30 kb, vorzugsweise größer als 40 kb, besonders bevorzugt größer als 60 kb sind, in heterologe Wirtszellen zu übertragen und dort eine Etablierung von Fremd-DNA zu gewährleisten, für eine Spinosyn-Produktion geeignet. Vorzugsweise werden solche BAC-, PAC- und funktionell gleichwertigen Vektoren verwendet, die zu einem Shuttle-Vektor modifiziert sind und z.B. eine Plasmidreplikation sowohl in Gram-negativen Bakterien, wie *Escherichia coli*, als auch in Gram-positiven Bakterien, wie *Streptomyces*, erlauben. Solche bevorzugten Shuttle-Vektoren können DNA-Fragmente einer Größe tragen, die in üblichen Vektoren, wie z.B. Cosmidvektoren, nicht klonierbar sind und nicht in heterologe Wirte, wie Actinomyceten, z.B. Streptomyceten, übertragbar sind. Letztere Vektoren können sowohl durch Transformation, Konjugation, Elektroporation, Protoplastentransformation oder andere geeignete Verfahren übertragen werden. In hervorragender Weise sind solche Shuttle-Vektoren innerhalb einer heterologen Population Gram-negativer oder Gram-positiver Bakterien, zwischen Gram-positiven und Gram-negativen Bakterien, zwischen Bakterien und Archea, zwischen Pro- und Eukaryonten konjugativ übertragbar. Die in heterologe Wirte, wie z.B. Streptomyceten, übertragenen BAC-, PAC- oder funktionell gleichwertigen Shuttle-

Vektoren können autonom repliziert werden oder ins Genom des Wirtes integriert werden. Letztere Integration kann über homologe Rekombination, über einen  $\Phi$ C31-Integrationsmechanismus (Hopwood et al., 1985), über ortsspezifische Integration, die von pSAM2 (Smokvina et al., 1990; WO 95/16046) determinierten Funktionen abhängt oder über Mini-Circle vermittelte Funktionen (Motamedi et al., 1995; WO 96/00282) erfolgen.

Solche Shuttle-Vektoren erlauben es, spezifische, durch außerordentlich große DNA-Bereiche determinierte Biosynthesewege von Primär- oder Sekundärmetaboliten, durch Transfer eines einzigen rekombinanten Vektors heterolog in besonders geeigneten Wirtszellen zu exprimieren. So kann das identifizierte Cluster für die Biosynthese von Spinosyn in Organismen, wie Actinomyceten, z.B. *Streptomyces*, durch Transfer eines einzigen rekombinanten Shuttle-Vektors ausgeprägt werden. Aufgrund der Größe dieses Biosyntheseclusters ist diese heterologe Expression der Spinosyn-Biosynthese mit einem einzigen Cosmidvektor nicht möglich. Die Übertragung eines rekombinanten BAC-, PAC- oder funktionell gleichwertigen Shuttle-Vektors, der die erfindungsgemäßen Nucleinsäuren trägt, kann zu einer signifikanten Steigerung der Produktion von Spinosyn im Vergleich zur Spinosyn-Produktion des Stammes *S. spinosa* oder abgeleiteter Mutanten mit erhöhter Spinosyn-Bildung führen. Zudem kann ein solcher, für die Spinosyn-Biosynthese codierender Shuttle-Vektor genutzt werden, um nach Übertragung in heterologe Wirtszellen deren Biosynthese- und Modifizierungsleistung auszunutzen, um zu einer signifikanten Modifizierung von Spinosyn oder Spinosyn-Biosynthesevorstufen zu gelangen. Hierdurch ist es zudem möglich, neue Spinosyn-Derivate durch den Transfer eines einzigen rekombinanten Vektors in heterologe Wirtszellen herzustellen.

Weiterhin können solche Shuttle-Vektoren verwendet werden: klonierte Biosynthesewege von Sekundärmetaboliten als Bestandteil eines einzigen rekombinanten Shuttle-Vektors genetisch zu modifizieren. Solche Modifizierungen können z.B. in einem *E. coli* Wirt durchgeführt werden, z.B. unter Ausnutzung von Rekombinationsereignissen unter Beteiligung des recA-Genproduktes oder der recE-

und recT-Genprodukte (Muyrers et al., 1999). Weiterhin können solche Vektoren durch *in vitro*-Verfahren, wie z.B. das Template Generation System (Finnzymes, FIN-02201, Espoo, Finnland) oder das Transposomics-System (Epicentre Technologies, Biozym Diagnostika GmbH, Oldendorf, Deutschland) modifiziert werden. Solche, für veränderte Biosynthesewege codierende Shuttle-Vektoren können dann in geeignete Wirtszellen übertragen werden, um zur Produktion veränderter Sekundärmetabolite zu gelangen. In analoger Weise können die genannten Shuttle-Vektoren genutzt werden, die erfindungsgemäßen Nucleinsäuren zu modifizieren, um sie dann, nach Transfer in geeignete Wirtszellen zur Produktion veränderter Spinosyne einzusetzen.

Teile der erfindungsgemäßen Nucleinsäuren können auch als Bestandteil von zwei oder mehreren Vektoren, wie z.B. Cosmidvektoren, eine genetische Information determinieren, die in Kombination miteinander zur Biosynthese von Spinosyn oder Spinosyn-Vorstufen, wie z.B. Pseudoaglycon oder Spinosyn-Aglycon geeignet sind. Solche Kombinationen von rekombinanten Vektoren können eingesetzt werden, um zu einer Spinosyn-Produktion in anderen Organismen als *S. spinosa* zu gelangen. Dies kann bei einer Expression in besonders geeigneten Wirten zu einer signifikanten Steigerung der Spinosyn-Produktion im Vergleich zu *S. spinosa* oder abgeleiteten produktionsverstärkten Mutanten führen. Weiterhin ist es möglich, erfindungsgemäße Nucleinsäuren in einzelnen rekombinanten Vektoren dieser Vektorkombination so zu verändern, dass eine heterologe Produktion von Spinosyn-Derivaten in Wirtszellen möglich ist. Desweiteren kann eine solche Kombination von rekombinanten Vektoren durch deren Transfer in heterologe Wirte zur Bildung neuer Spinosyn-Derivate unter Ausnutzung des wirtseigenen Enzymsystems geeignet sein.

Gegenstand der vorliegenden Erfindung sind auch Wirtszellen, die zumindest eine der erfindungsgemäßen Nucleinsäuren enthalten. Als Wirtszelle eignen sich sowohl prokaryotische Zellen, vorzugsweise Actinomyceten, besonders bevorzugt Streptomyceten, als auch eukaryotische Zellen, wie Säugerzellen, Pflanzenzellen oder Hefezellen.

In besonderer Weise können die erfindungsgemäßen Nucleinsäuren in pflanzliche Zellen übertragen und exprimiert werden. Hierdurch können transgene Pflanzen hergestellt werden, die das pflanzenschützende, insektizide Spinosyn bzw. Derivate davon produzieren. Eine Übertragung der erfindungsgemäßen Nucleinsäuren in die Pflanzenzellen oder pflanzliche Zellkulturen kann mit üblichen Verfahren u.a. auch durch Partikelbeschuss erfolgen.

Gegenstand der vorliegenden Erfindung sind weiterhin die Polypeptide, die von den erfindungsgemäßen Nucleinsäuren codiert werden. Die erfindungsgemäßen Polypeptide können ein vollständiges Enzym darstellen, das einen Schritt der Spinosyn-Biosynthese katalysiert. Jedoch sind auch solche Polypeptide von der Erfindung erfasst, die nur einen Teil der vollständigen Aminosäuresequenz des betreffenden Enzyms aufweisen.

Der Ausdruck "Teilsequenz", wie er hierin verwendet wird, bezieht sich somit auf die Aminosäuresequenz eines Polypeptids, das noch die Aktivität des entsprechenden vollständigen Enzyms oder einer enzymatisch aktiven Domäne ausüben kann.

Im Folgenden werden bevorzugte erfindungsgemäße Nucleinsäuren und Polypeptide mit Bezug auf die entsprechenden SEQ ID NOS näher charakterisiert.

SEQ ID NOS: 7 und 8, ORF1:

Nucleotidpositon 828 bis 1 der SEQ ID NO: 4, 275 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

SEQ ID NOS: 9 und 10, ORF2:

Nucleotidposition 1.283 bis 2.455 der SEQ ID NO: 4, 390 Aminosäuren;  
das ableitbare Genprodukt ist eine Glycosyltransferase.

SEQ ID NOS: 11 und 12, ORF3:

Nucleotidposition 2.495 bis 3.247 der SEQ ID NO: 4, 250 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

5      SEQ ID NOS: 13 und 14, ORF4:

Nucleotidposition 4.440 bis 3.253 der SEQ ID NO: 4, 395 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

SEQ ID NOS: 15 und 16, ORF5:

10      Nucleotidposition 4.578 bis 6.197 der SEQ ID NO: 4, 539 Aminosäuren;  
das ableitbare Genprodukt ist ein C-C verknüpfendes Enzym, das Cyclisierungs-  
reaktionen durchführt.

SEQ ID NOS: 17 und 18, ORF6:

15      Nucleotidposition 6.211 bis 7.404 der SEQ ID NO: 4, 397 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

SEQ ID NOS: 19 und 20, ORF7:

20      Nucleotidposition 7.401 bis 8.300 der SEQ ID NO: 4, 299 Aminosäuren;  
das ableitbare Genprodukt ist eine Methyltransferase.

SEQ ID NOS: 21 und 22, ORF8:

Nucleotidposition 8.300 bis 9.466 der SEQ ID NO: 4, 388 Aminosäuren;  
das ableitbare Genprodukt ist ein Enzym, das an Cyclisierungsreaktionen beteiligt ist.

25

SEQ ID NOS: 23 und 24, ORF9:

Nucleotidposition 10.572 bis 9.562 der SEQ ID NO: 4, 336 Aminosäuren.  
das ableitbare Genprodukt ist eine 2,3-Reduktase.

SEQ ID NOS: 25 und 26, ORF10:

Nucleotidposition 12.029 bis 10.569 der SEQ ID NO: 4, 486 Aminosäuren;  
das ableitbare Genprodukt ist eine 2,3-Dehydratase.

5 SEQ ID NOS: 27 und 28, ORF11:

Nucleotidposition 12.549 bis 12.109 der SEQ ID NO: 4, 146 Aminosäuren; das  
ableitbare Genprodukt hat Homologien zu einer Thioesterase.

10 SEQ ID NOS: 29 und 30, ORF12:

Nucleotidposition 13.865 bis 12.546 der SEQ ID NO: 4, 439 Aminosäuren;  
das ableitbare Genprodukt ist eine Glykosyltransferase.

15 SEQ ID NOS: 31 und 32, ORF13:

Nucleotidposition 14.245 bis 15.633 der SEQ ID NO: 4, 462 Aminosäuren;  
das ableitbare Genprodukt ist eine 3,4-Dehydratase.

20 SEQ ID NOS: 33 und 34, ORF14:

Nucleotidposition 15.671 bis 16828 der SEQ ID NO: 4, 385 Aminosäuren;  
das ableitbare Genprodukt ist eine 4-Aminotransferase.

25 SEQ ID NOS: 35 und 36, ORF15:

Nucleotidposition 16.831 bis 17.580 der SEQ ID NO: 4, 249 Aminosäuren;  
das ableitbare Genprodukt ist eine N-Dimethyltransferase.

30 SEQ ID NOS: 37 und 38, ORF16:

Nucleotidposition 18.930 bis 18.205 der SEQ ID NO: 4, 241 Aminosäuren;  
das ableitbare Genprodukt ist eine 3,4-Reduktase.

SEQ ID NOS: 39 und 40, ORF17:

Nucleotidposition 19.025-19.861 der SEQ ID NO: 4, 278 Aminosäuren;  
das ableitbare Genprodukt ist ein Transkriptions-Regulator.

SEQ ID NOS: 41 und 42, ORF18:

Nucleotidpositionen 116-7903 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 2595:

- 5 Nucleotidpositionen 128-1402, Aminosäure Positionen 5-429 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;  
Nucleotidpositionen 1691-2656, Aminosäurepositionen 526-847 codieren eine Acyltransferase-Domäne;  
Nucleotidpositionen 2798-3052, Aminosäurepositionen 895-979 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;  
10 Nucleotidpositionen 3107-4372, Aminosäurepositionen 998-1419 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;  
Nucleotidpositionen 4688-5662, Aminosäurepositionen 1525-1849 codieren eine Acyltransferase-Domäne;  
15 Nucleotidpositionen 6587-7138, Aminosäurepositionen 2158-2341 codieren eine Ketoreduktase-Domäne;  
Nucleotidpositionen 7409-7666, Aminosäurepositionen 2432-2517 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

20 SEQ ID NOS: 43 und 44, ORF19:

Nucleotidpositionen 7921-14379 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 2152:

- Nucleotidpositionen 8029-9318, Aminosäurepositionen 37-466 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;  
25 Nucleotidpositionen 9634-10608, Aminosäurepositionen 572-896 codieren eine Acyltransferase-Domäne;  
Nucleotidpositionen 10705-11259, Aminosäurepositionen 929-1113 codieren eine Dehydratase-Domäne;  
Nucleotidpositionen 12043-13080, Aminosäurepositionen 1375-1720 codieren eine  
30 Enoylreduktase-Domäne;

Nucleotidpositionen 13093-13635, Aminosäurepositionen 1725-1905 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 13885-14142, Aminosäurepositionen 1989-2074 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

5

SEQ ID NOS: 45 und 46, ORF20:

Nucleotidpositionen 14424-23936 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 3170:

10

Nucleotidpositionen 14523-15824, Aminosäurepositionen 34-467 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

Nucleotidpositionen 16110-17075, Aminosäurepositionen 563-884 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 17997-18536, Aminosäurepositionen 1192-1371 codieren eine Ketoreduktase-Domäne;

15

Nucleotidpositionen 18795-19052, Aminosäurepositionen 1458-1543 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

Nucleotidpositionen 19107-20387, Aminosäurepositionen 1562-1988 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

20

Nucleotidpositionen 20718-21692, Aminosäurepositionen 2099-2423 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 22620-23171, Aminosäurepositionen 2733-2916 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 23436-23693, Aminosäurepositionen 3005-3090 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

25

SEQ ID NOS: 47 und 48, ORF21:

Nucleotidpositionen 23983-38757 der SEQ ID NO: 5, Aminosäurepositionen 1 bis 4924:

30

Nucleotidpositionen 24082-25392, Aminosäurepositionen 34-470 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;



Nucleotidpositionen 25696-26661, Aminosäurepositionen 572-893 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 26761-27315, Aminosäurepositionen 927-1111 codieren eine Dehydratase-Domäne;

5 Nucleotidpositionen 28231-28782, Aminosäurepositionen 1417-1600 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 29035-29265, Aminosäurepositionen 1685-1761 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

10 Nucleotidpositionen 29329-30624, Aminosäurepositionen 1783-2214 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

Nucleotidpositionen 30928-31902, Aminosäurepositionen 2316-2640 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 32827-33378, Aminosäurepositionen 2949-3132 codieren eine Ketoreduktase-Domäne;

15 Nucleotidpositionen 33652-33900, Aminosäurepositionen 3224-3306 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;

Nucleotidpositionen 33952-35262, Aminosäurepositionen 3324-3760 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

20 Nucleotidpositionen 35554-36522, Aminosäurepositionen 3858-4180 codieren eine Acyltransferase-Domäne;

Nucleotidpositionen 37453-37998, Aminosäurepositionen 4491-4672 codieren eine Ketoreduktase-Domäne;

Nucleotidpositionen 38254-38511, Aminosäurepositionen 4758-4843 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne.

25

SEQ ID NOS: 49 und 50, ORF22:

Nucleotidpositionen 38808-50000 der SEQ ID NO: 5 und die Nukleotidpositionen 1 bis 5574 der SEQ ID NO: 6, Aminosäurepositionen 1 bis 5588:

30 Nucleotidpositionen 38907-40226 der SEQ ID NO: 5, Aminosäurepositionen 34-473 codiert eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;

- Nucleotidpositionen 40494-41453 der SEQ ID NO: 5, Aminosäurepositionen 563-882 codieren eine Acyltransferase-Domäne;
- Nucleotidpositionen 41556-42119 der SEQ ID NO: 5, Aminosäurepositionen 917-1104 codieren eine Dehydratase-Domäne;
- 5 Nucleotidpositionen 43017-43568 der SEQ ID NO: 5, Aminosäurepositionen 1404-1587 codieren eine Ketoreduktase-Domäne;
- Nucleotidpositionen 43833-44090 der SEQ ID NO: 5, Aminosäurepositionen 1676-1761 codieren eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein Domäne;
- 10 Nucleotidpositionen 44151-45473 der SEQ ID NO: 5, Aminosäurepositionen 1782-2222 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;
- Nucleotidpositionen 45765-46730 der SEQ ID NO: 5, Aminosäurepositionen 2320-2641 codieren eine Acyltransferase-Domäne;
- Nucleotidpositionen 46827-47459 der SEQ ID NO: 5, Aminosäurepositionen 2674-2884 codieren eine Dehydratase-Domäne;
- 15 Nucleotidpositionen 48378-48935 der SEQ ID NO: 5, Aminosäurepositionen 3191-3376 codieren eine Ketoreduktase-Domäne;
- Nucleotidpositionen 49182-49412 der SEQ ID NO: 5, Aminosäurepositionen 3459-3535 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;
- 20 Nucleotidpositionen 49482-50000 der SEQ ID NO: 5 und Nucleotidposition 1 bis 759 der SEQ ID NO: 6, Aminosäurepositionen 3559-3984 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne;
- Nucleotidpositionen 1084-2049 der SEQ ID NO: 6, Aminosäurepositionen 4093-4414 codieren eine Acyltransferase-Domäne;
- 25 Nucleotidpositionen 2146-2697 der SEQ ID NO: 6, Aminosäurepositionen 4447-4630 codieren eine Dehydratase-Domäne;
- Nucleotidpositionen 3604-4155 der SEQ ID NO: 6, Aminosäurepositionen 4933-5116 codieren eine Ketoreduktase-Domäne;
- Nucleotidpositionen 4420-4677 der SEQ ID NO: 6, Aminosäurepositionen 5205-5290 codieren eine  $\beta$ -Ketoacyl:Acyl Carrier Protein-Domäne;
- 30 Nucleotidpositionen 4864-5538 der SEQ ID NO: 6, Aminosäurepositionen 5353-5577 codieren eine Thioesterase-Domäne.

SEQ ID NOS: 52 und 53, ORF23:

Nucleotidposition 344 bis 1333 der SEQ ID NO: 51, 329 Aminosäuren;  
das ableitbare Genprodukt ist eine dNDP-Glucose-4,6-Dehydratase.

5

SEQ ID NOS: 54 und 55, ORF24:

Nucleotidposition 1330 bis 2247 der SEQ ID NO: 51, 305 Aminosäuren;  
das ableitbare Genprodukt ist eine dNDP-4-keto-6-Deoxyglucose-3,5-Epimerase.

- 10 Die an der Cyclisierung des 5, 6, 5- Tricyclus beteiligten Produkte des ORF 5 (SEQ  
ID NO: 16) und des ORF 8 (SEQ ID NO: 22) sind aufgrund der ungewöhnlichen  
Cyclisierungsreaktionen von besonderem Interesse. Daher beinhaltet die vorliegende  
Erfindung insbesondere auch homologe Nucleinsäuren oder homologe Genprodukte.  
Vorzugsweise zeigen diese homologen Genprodukte mindestens eine 50 %ige,  
15 bevorzugt eine 60 %ige und besonders bevorzugt eine 70 %ige Identität auf Amino-  
säureebene.

- Weiterhin sind Antikörper Gegenstand der Erfindung, die spezifisch an die  
vorstehend genannten Polypeptide binden. Die Herstellung solcher Antikörper erfolgt  
20 auf die übliche Weise. Diese Antikörper können genutzt werden, um Expres-  
sionsklone z.B. einer Genbank zu identifizieren, die die erfindungsgemäßen Nuclein-  
säuren tragen.

- Gegenstand der vorliegenden Erfindung sind auch Verfahren zum Herstellen der  
25 erfindungsgemäßen Nucleinsäuren. Die erfindungsgemäßen Nucleinsäuren können  
auf die übliche Weise hergestellt werden. Beispielsweise können die Nucleinsäure-  
moleküle vollständig chemisch synthetisiert werden. Man kann auch kurze Stücke  
der erfindungsgemäßen Nucleinsäuren chemisch synthetisieren und solche  
Oligonucleotide radioaktiv oder mit einem Fluoreszenzfarbstoff markieren. Die  
30 markierten Oligonucleotide können auch verwendet werden, um Genbanken von  
Organismen zu durchsuchen. Klone, an die die markierten Oligonucleotide hybridi-

sieren, werden zur Isolierung der betreffenden DNA ausgewählt. Nach der Charakterisierung der isolierten DNA erhält man auf einfache Weise die erfindungsgemäßen Nucleinsäuren. Die erfindungsgemäßen Nucleinsäuren können auch mittels PCR-Verfahren unter Verwendung chemisch synthetischer Oligonucleotide hergestellt werden.

Gegenstand der vorliegenden Erfindung sind weiterhin Verfahren zum Herstellen der erfindungsgemäßen Polypeptide. Zur Herstellung der Polypeptide, die von den erfindungsgemäßen Nucleinsäuren codiert werden, können Wirtszellen, die zumindest eine der erfindungsgemäßen Nucleinsäuren enthalten, unter geeigneten Bedingungen kultiviert werden. Die gewünschten Polypeptide können danach auf übliche Weise aus den Zellen oder dem Kulturmedium isoliert werden. Die Polypeptide können auch in *in vitro*-Systemen hergestellt werden.

Das isolierte und charakterisierte Gencluster und benachbarte oder assoziierte DNA-Regionen stellen ein Target zur Steigerung der Spinosyn-Biosynthese durch genetische Manipulation, Über- oder Unterexpression von direkt oder indirekt an der Biosynthese involvierten Genen oder regulatorischen Sequenzen dar. Diese Manipulationen können sowohl in natürlichen Spinosyn-produzierenden Organismen als auch in gentechnisch hergestellten Spinosyn-produzierenden Organismen durchgeführt werden. Beispielsweise können ausgewählte ORF's unter die Kontrolle üblicher starker Promotoren wie dem *mel*-Promotor des Plasmides pIJ702 (John Innes Foundation, Norwich, UK, 1985) gestellt werden.

Durch die Klonierung und Identifizierung der Spinosyn-Biosynthesegene schafft die vorliegende Erfindung die genetische Basis, mittels molekulargenetischer Verfahren neue Spinosyn-Vorstufen und -Derivate herzustellen.

Der Ausdruck "Spinosyn-Vorstufen", wie er hierin verwendet wird, bezieht sich auf alle nachweisbaren und alle postulierbaren Biosynthesevorstufen von Spinosyn.

Der Ausdruck "Spinosyn-Derivate", wie er hierin verwendet wird, bezieht sich auf Strukturderivate aller bisher bekannten Spinosyne.

Gegenstand der vorliegenden Erfindung ist somit auch ein Verfahren zum Herstellen von Spinosyn-Vorstufen und -Derivaten.

Die erfindungsgemäßen Nucleinsäuren können beispielsweise eingesetzt werden, um durch kombinatorische Biosynthese neue Spinosyn-Derivate mit Veränderungen des Spinosyn-Aglycons herzustellen. Dies kann z.B. dadurch erreicht werden, dass die von ORF 19 codierte, eine Acetat-Einheit einbauende Acyltransferase-Domäne ausgetauscht wird gegen eine Acyltransferase-Domäne, die eine Propionat-Einheit einbaut. In gleicher Weise kann die, eine Acetat-Einheit einbauende Acyltransferase-Domäne des ORF 18 gegen eine Acyltransferase-Domäne ausgetauscht werden, die eine Propionat-Einheit einbaut. Ferner ist es möglich beide oder jeweils eine Ketoreduktase-Domäne, die von beiden genannten ORF's codiert werden zu inaktivieren, durch eine inaktive Ketoreduktase-Domäne zu ersetzen oder zu deletieren, wodurch eine Hydroxygruppe an der entsprechenden Position im Makrocyclus biosynthetisch hergestellt werden kann. Alle Acyltransferase-, Ketoreduktase-, Dehydratase-, Enoylreduktase-,  $\beta$ -Ketoacyl:Acyl Carrier Protein und Thioesterase-Domänen können einzeln oder in beliebiger Kombination durch entsprechende Polyketidsynthase-Domänen mit anderer Substrat- oder Reaktionsspezifität ersetzt werden, in beliebiger Kombination miteinander fusioniert, einzeln oder in beliebiger Kombination mutagenisiert, deletiert oder dupliziert werden. Ferner können Modulcodierende Sequenzen ausgetauscht werden. So ist es denkbar die Modul 2-codierende DNA-Sequenz (Abb. 6) gegen die Modul 1- oder Modul 3, 4, 5, 6, 7, 8- oder Modul 9-codierende DNA-Sequenz (Abb. 6) zu ersetzen und funktionell zu exprimieren. Es ist auch denkbar die Modul 2-codierende DNA-Sequenz oder jede andere Modul-codierende DNA-Sequenz des Spinosyn-Polyketidsynthase-Genclusters gegen eine andere Modul-codierende DNA-Sequenz des Spinosyn-Polyketidsynthase-Genclusters, die eine andere biosynthetische Verlängerungseinheit einbaut, auszutauschen. Darüber hinaus kann jede andere Modul-codierende DNA-

Sequenz des Spinosyn-Polyketidsynthase-Genclusters gegen eine andere Modul-codierende DNA-Sequenz einer anderen Polyketidsynthase-Nukleinsäuresequenz aus *S. spinosa* oder einem anderen Organismus als *S. spinosa*, wie z.B. *Saccharopolyspora erythraea*, ausgetauscht werden. Diese Veränderungen können  
5 unter Ausnutzung der ET-Rekombination (WO 99/29837; Muyrers et al., 1999) oder anderer Klonierungs- und Rekombinationstechniken durchgeführt werden.

Gegenstand der Erfindung sind somit auch alle Modul- oder Domänen-codierenden Nucleinsäuren, die natürlicher oder gentechnisch erzeugter Bestandteil der Spinosyn-Polyketidsynthase sind.  
10

Der Ausdruck "Modul", sowie er hierin verwendet wird, bedeutet, dass eine Anordnung von drei enzymkatalytisch aktiven Domänen vorliegt, die zu eine Verlängerung der wachsenden Polyketidkette um eine biosynthetische Verlängerungseinheit  
15 führen. Bei diesen Domänen handelt es sich um eine  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne, eine Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein-Domäne. Ein Modul kann auch eine Ketoreduktase-, eine Dehydratase-, eine Enoylreduktase- und eine Thioesterase-Domäne tragen. Ein sog. Ladungsmodul, das am Beginn der Biosynthese steht kann von den genannten Domänen lediglich eine  
20 Acyltransferase-Domäne und eine  $\beta$ -Ketoacyl:Acyl-Carrier Protein-Domäne tragen, sowie eine enzymatisch inaktive  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase-Domäne. Eine Polyketidsynthase-Domäne umfasst jeweils eine dieser genannten enzymatischen Aktivitäten.

Die erfindungsgemäßen Nucleinsäuren können weiterhin genutzt werden, um im Zuge einer kombinatorischen Biosynthese durch Neuordnung und Expression von Spinosyn-Polyketidsynthase-Nucleinsäuresequenzen oder durch Kombination und Expression zusammen mit Polyketidsynthase-Nucleinsäuresequenzen einer anderen Polyketidsynthase codierenden Nucleinsäuresequenz aus *S. spinosa* oder einem  
30 anderen Organismus, wie z.B. *Saccharopolyspora erythraea*, Bibliotheken von rekombinanten Polyketidsynthase-Nucleinsäuresequenzen, rekombinanten Polyketid-

synthase-Proteinen oder rekombinant erzeugten Polyketiden herzustellen. Diese Polyketide können durch die Verwendung der erfindungsgemäßen Nucleinsäuren oder die Verwendung anderer Nucleinsäuren, deren ableitbaren Produkte an der Biosynthese anderer Zucker und Ankopplung ans Aglycon beteiligt sind, glycosyliert werden. Es ist bekannt, dass die Glycosylierung des Aglycons eine entscheidende Rolle bei der biologischen Aktivität am Wirkort spielt. Diese Veränderungen können sowohl in natürlichen als auch in gentechnisch hergestellten Spinosyn-produzierenden Organismen, insbesondere Bakterien, erfolgen. Weiterhin können diese Veränderungen unter Ausnutzung der ET-Rekombination (WO 99/29837; Muyrers et al., 1999) oder anderer Klonierungs- und Rekombinationstechniken durchgeführt werden.

Die erfindungsgemäßen Nucleinsäuren, Vektoren und regulatorischen oder genetisch mobilen Regionen können außerdem zum Auffinden von Genen verwendet werden, die für Polypeptide codieren, welche funktionell ähnliche Polyketidsynthasen oder funktionell ähnliche Produkte, die an einer Zuckerbiosynthese beteiligt sind, codieren.

Da die erfindungsgemäßen Nucleinsäuren einen umfangreichen Teil des Genoms von *S. spinosa* ausmacht, können die erfindungsgemäßen Nucleinsäuren als Marker bei der Sequenzierung des Genoms von *S. spinosa* eingesetzt werden, wodurch die Anordnung von Teilsequenzen eines Genomsequenzierungsprojektes erheblich erleichtert wird.

Somit liefern die erfindungsgemäßen Nucleinsäuren Daten, die im Rahmen eines Genomsequenzierungsprojektes und eines sich darauf aufbauenden Metabolic Engineering zur Steigerung der Spinosynproduktion eingesetzt werden können.

**Erläuterungen zu den Abbildungen:**

Abbildung 1: Modell für die Biosynthese der Spinosyn-Zucker D-Forosamin und 2, 3, 4-Tri-O-Methyl-L-Rhamnose.

5

Abbildung 2: Lage, der an der Spinosyn-Biosynthese direkt oder indirekt beteiligten DNA-Regionen 1 (SEQ ID NO: 4) und DNA-Region 2 (SEQ ID NOS: 5 und 6). Die schwarzen Balken im unteren Teil der Abbildung geben schematisch die Positionen der Cosmid-DNA Inserts zueinander und in Bezug zu den DNA-Regionen 1 und 2 an. Die dargestellten Cosmid-Inserts wurden zur Sequenzierung der SEQ ID NOS: 1 bis 3 herangezogen.

10

Abbildung 3: Schematische Darstellung der Lage der Insert-DNA (schwarze Balken im unteren Teil der Abbildung) der benannten Cosmide, die zur Ankopplung eines Forosamin-Restes oder eines Trimethyl-Rhamnose-Restes durch Biotransformation des Spinosyn-Aglycons und Spinosyn-Pseudoaglycons herangezogen worden sind.

15

Abbildung 4: Schematische Darstellung der offenen Leserahmen (ORF's) der DNA-Region 3, die der SEQ ID NO: 51 entspricht, auf Cosmid 16-2-2.

20

Abbildung 5: Schematische Darstellung offener Leserahmen (ORF's) der DNA-Regionen 1 und 2. Die ORF's sind nummeriert von 1 bis 22 entsprechend SEQ ID NOS: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47 und 49.

25

Abbildung 6: Schematische Darstellung offener Leserahmen (ORF's) der DNA-Region 2 (SEQ ID NOS: 5 und 6) und ableitbarer Module und Domänen. SM, Startmodul; M1 bis M10, Modul 1 bis Modul 10; KS,  $\beta$ -Ketoacyl:Acyl Carrier Protein Synthase; AT, Acyltransferase; ACP,  $\beta$ -Ketoacyl:Acyl Carrier Protein; KR, Ketoreduktase; DH, Dehydratase; ER, Enoylreduktase.

30



Abbildung 7: Schematische Darstellung der Lage von BAC-Shuttleklon Insert-DNA als schwarze Balken im unteren Teil der Abbildung. Die Größe der Insert-DNA beträgt mindestens 100 kb. Durchgezogene Balken: DNA-Sequenz ist identisch mit Teilen bzw. der Gesamtheit der DNA-Regionen 1 und 2. Gestrichelte Balken: DNA-Sequenz liegt außerhalb des sequenzierten Bereichs.

5

## Beispiele

### Bakterienstämme und Plasmide

5 *Escherichia coli* XL1-Blue MRF' und die Cosmidvektoren SuperCos1 (Stratagene, Europe) und pOJ446 (Biermann et al., 1992) wurden verwendet zur Erstellung von Genbanken von *S. spinosa* ATCC49460 (American Type Culture Collection, U.S.A., EP-A 0 375 316). *E. coli* JM110 (Stratagene, Europe) wurde verwendet zur Propagierung von Plasmiden, die durch Transformation nach *Streptomyces* übertragen wurden. *Streptomyces albus* J1074 (Chater and Wilde, 1980; John Innes Institut in 10 Norwich, UK) wurde zur heterologen Expression von und zur Biotransformation mit Spinosyn-Biosynthesegenen eingesetzt.

15 Plasmid pBeloBAC11 (Kim et al., 1996) und pOJ446 (Biermann et al., 1992) wurden verwendet zur Herstellung eines *E. coli* – *Streptomyces* BAC-Shuttlevektors.

### Molekularbiologische Methoden

20 Molekularbiologische Methoden wie DNA-Restriktion, Agarose-Gelelektrophorese von DNA, Ligation von Restriktionsfragmenten, Kultivierung und Transformation von *E. coli* wurden durchgeführt wie beschrieben in Sambrook et. al (1989). Plasmide wurden mit Qiagen Plasmid Kit (Qiagen, Hilden, Deutschland) isoliert. Die verwendeten Enzyme stammten von Roche Diagnostics GmbH (Mannheim, Deutschland).

25 Anzucht Bedingungen and molekulargenetische Methoden mit *S. spinosa* und Streptomyceten sind beschrieben in (Hopwood et al., 1985). Alle Anzuchten in Flüssigkultur von *S. spinosa* oder Streptomyceten erfolgten aerob in Erlenmeyerkolben bei 28°C.

Die DNA-DNA-Hybridisierungen erfolgten unter Verwendung des DIG-High Prime DNA Labeling and Detection Kit nach Angaben des Herstellers (Roche Diagnostics GmbH, Mannheim, Deutschland).

5 Wachstumsmedien:

LB Sambrook et. al., 1989

TS Difco Bestell-Nummer 0 370-17-3 (Difco Detroit, MI, USA)

10

R5A Illing et al., 1985

#### Herstellung einer Cosmid Genbank von *S. spinosa*

15 Um eine Genbank von *S. spinosa* zu erhalten, wurde chromosomale DNA von *S. spinosa* ATCC49460 mit *Mbo*I partiell geschnitten und durch Zentrifugation im Glucosedichtegradienten aufgetrennt. Die Cosmid-DNA (SuperCos1, Stratagene Europe) wurde nach Angaben des Herstellers vorbereitet, mit den *S. spinosa* DNA-Fragmenten zwischen 35 und 45 kb ligiert und mit Hilfe des Gigapack-Verpackungssystem (Stratagene Europe) in Phagenpartikel verpackt. Die Transfektion erfolgte in *E. coli* XL-1 blue MRF'. Diese Methode wurde ebenfalls dazu verwendet eine zweite *S. spinosa* Genbank anzulegen unter Verwendung des *E. coli-Streptomyces* Shuttle-Cosmids pOJ446.

20

25 Sequenzierung des Spinosynbiosynthese-Genclusters und eines DNA-Fragmentes das außerhalb dieses Clusters liegt, dessen Produkte aber an der Biosynthese von Spinosyn beteiligt sind

30

Die Insert-DNA der SuperCos1 Cosmide 16-1-8, 16-59-1, und 16-59-8 wurden sequenziert. Eine ca. 4 kb große Lücke zwischen den Cosmiden 16-59-1 und 16-1-8

wurde durch primer walking Sequenzierung eines entsprechenden Teilbereiches von Cosmid 16-59-6 geschlossen.

5 Eine ca. 2,3 kb große DNA-Sequenz auf dem SuperCos1 Cosmid 16-2-2 wurde sequenziert.

Identifizierung und Charakterisierung von chromosomalen DNA-Fragmenten einer BAC-Shuttlevektor-Genbank aus *S. spinosa*, die Spinosyn-Biosyntheseensequenzen tragen

10

Zur Herstellung des BAC-Shuttlevektors, der nicht nur in *E. coli* sondern auch in Actinomyceten wie *Streptomyces* übertragen und vermehrt werden kann, wurde der Vektor pBeloBAC11 mit *Xho*I linearisiert, und durch die Anwendung von Klenow Polymerase wurden glatte DNA-Enden hergestellt. Ein ca. 6 kb großes *Dra*I –  
15 *Eco*RV DNA-Fragment des Cosmidvektors pOJ446, das den Replikationsursprung des Plasmides SCP2\*, das Apramycinresistenzgen sowie den *oriT* zum konjugativen Transfer trägt, wurde mit dem linearisierten BAC-Vektor ligiert. Der resultierende Vektor erhielt die Bezeichnung pEBZ333.

20

Ausgehend von partiell mit *Mbo*I geschnittener genomischer DNA des Stammes *S. spinosa* ATCC49460 und dem mit *Bam*HI geschnittenen Vektor pEBZ333 wurde eine BAC-Genbank erstellt.

Analyse und Annotation offener Leserahmen direkt oder indirekt an der Spinosyn-Biosynthese beteiligter DNA-Sequenzen

25

Ausgehend von der Sequenz gemäß SEQ ID NOS: 1 bis 3 wurden offene Leserahmen (ORF's) identifiziert, die direkt oder indirekt an der Biosynthese von Spinosyn beteiligt sind. Diese ORF's wurden in zwei DNA-Regionen unterteilt, die  
30 als DNA-Region 1 und DNA-Region 2 (Abb. 2 und 5) bezeichnet werden und die

Sequenzen gemäß SEQ ID NO: 4 bzw. 5 und 6 tragen. Die DNA-Region 1 trägt offene Leserahmen, deren Produkte an der Modifizierung und Tricyclusbildung des Spinosyn-Aglycons beteiligt sind, während die DNA-Region 2 (Abb. 2, 5 und 6) offene Leserahmen umfasst, deren Produkte die Spinosyn-Polyketidsynthase codieren. Die beiden jeweils ersten Nucleotide dieser DNA-Regionen liegen unmittelbar nebeneinander (Abb. 2, 3 und 5).

Eine weitere DNA-Region 3 (SEQ ID NO: 51) liegt außerhalb dieses Clusters von DNA-Sequenzen und trägt offene Leserahmen, deren Produkte ebenfalls an der Biosynthese des Spinosyn-Zuckers Trimethyl-Rhamnose beteiligt sind.

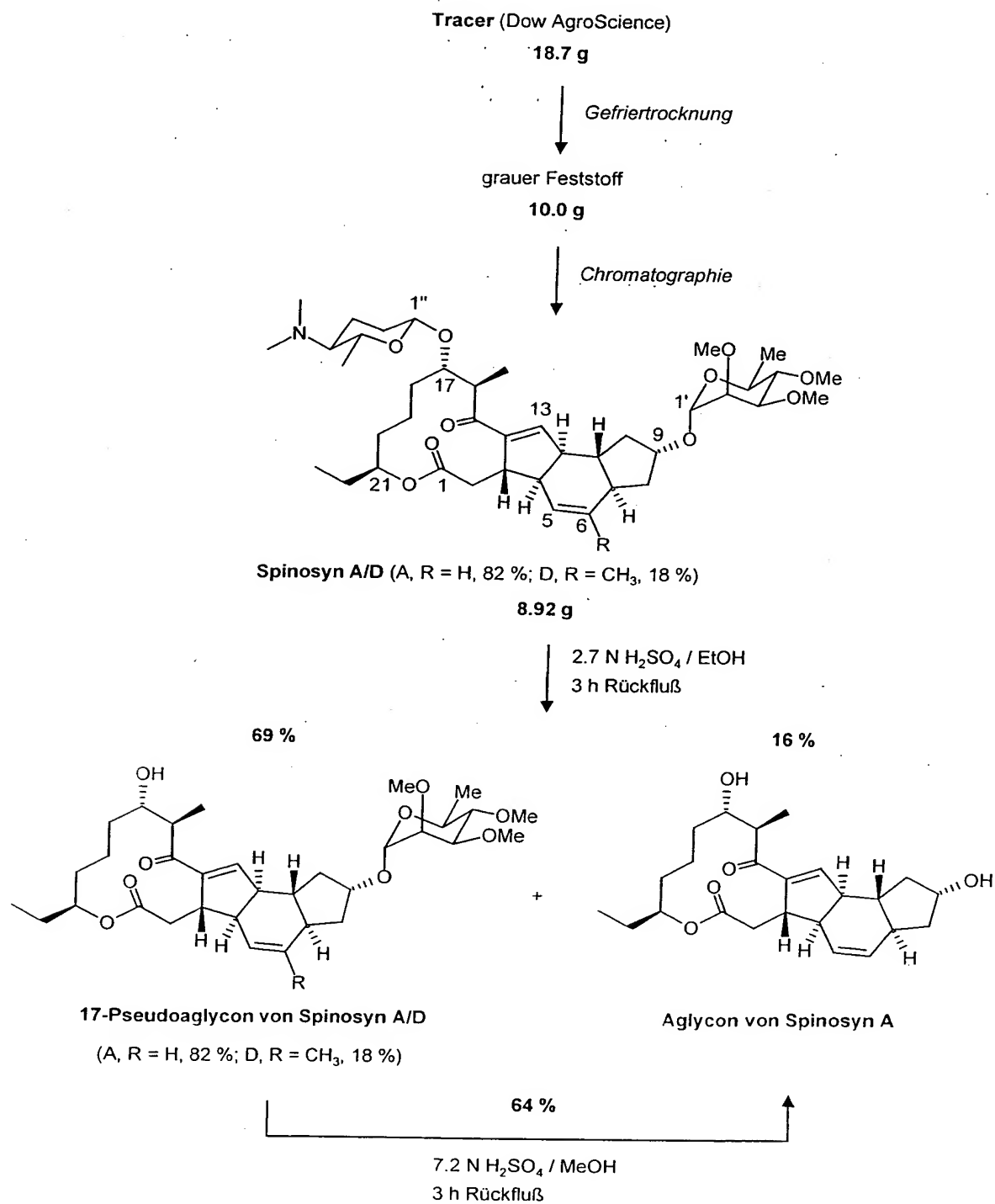
#### Herstellung des Spinosyn-Aglycons und 17-Pseudoaglycons aus Tracer®

Ausgehend von 18,7 g des kommerziell erhältlichen Produktes Tracer® wurden nach Gefriertrocknung und Säulenchromatographie an Kieselgel 8,92 g Spinosyn A und D in einem Verhältnis von 82:18 gewonnen.

Die Hydrolyse des Aminosuckers Forosamin gelang mit 2.7 N Schwefelsäure in Ethanol unter Rückfluß. Dabei fiel der Großteil des entstehenden 17-Pseudoaglycons von Spinosyn A/D aus. Im Filtrat wurden neben weiterem 17-Pseudoaglycon in Abhängigkeit von der Reaktionsdauer geringe bis mittlere Mengen des Spinosyn-Aglycons gefunden.

Eine vollständige Hydrolyse zum Aglycon gelang unter etwas drastischeren Bedingungen (7.2 N Schwefelsäure in Methanol unter Rückfluß). Die Aglycon-Fraktion enthielt ausschließlich Aglycon von Spinosyn A. Dies stimmt sehr gut mit der Literatur überein (Creemer et al., 1998), die unter entsprechenden Reaktionsbedingungen eine vollständige Zersetzung des Pseudoaglycons von Spinosyn D beschreibt. Als Ursache nehmen die Autoren eine leichtere Protonierung der 5,6-Doppelbindung bei Spinosyn D unter Bildung eines tertiären Carbokations und anschließende Umlagerungen an.

Es konnten somit ausgehend von 18,7 g käuflichem Tracer® 3,0 g Aglycon von Spinosyn A hergestellt werden.



### Gewinnung von Spinosyn A/D aus Tracer®

Die Gefriertrocknung von 18,7 g Tracer® lieferten 10,0 g grauen Feststoff. Nach Säulenchromatographie dieses Feststoffes an 800 cm<sup>3</sup> Kieselgel (Eluent: Dichlormethan/Methanol 95:5) erhielt man 8,92 g reines Spinosyn A/D (82 % A, 18 % D).  
5 – DC:  $R_f$  (SiO<sub>2</sub>, Dichlormethan/Methanol 9:1) = 0,46. – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta$  = 6,77 (s, 13-H); 5,88 (d, 5-H von Spinosyn A); 5,80 (m, 6-H von Spinosyn A); 5,49 (m, 5-H von Spinosyn D); 4,87 (d, 1'-H); 4,67 (m, 21-H); 4,43 (d, 1''-H); 4,31 (m, 9-H) u. a. – LC/MS: Elektrospray Positiv; Peak bei RT 44,0 min:  $m/z$  = 733 (100 %) [M+H]<sup>+</sup> (Spinosyn A); Peak bei 44,7 min:  $m/z$  = 747 (100 %) [M+H]<sup>+</sup> (Spinosyn D).  
10

### Darstellung des 17-Pseudoaglycons von Spinosyn A/D:

8,65 g (11,81 mmol) Spinosyn A/D wurden in 61 ml Ethanol gelöst und mit 104 ml  
15 Wasser und 208 ml 4 N H<sub>2</sub>SO<sub>4</sub> versetzt. Nach 3 h Erwärmen unter Rückfluß wurde der ausgefallene Feststoff (A) abfiltriert und getrennt vom Filtrat (B) aufgearbeitet. Der Feststoff (A) wurde mit 1 N H<sub>2</sub>SO<sub>4</sub> gewaschen, in 140 ml Dichlormethan aufgenommen, nacheinander mit gesätt. NaHCO<sub>3</sub>-Lösung und gesätt. NaCl-Lösung gewaschen, über Na<sub>2</sub>SO<sub>4</sub> getrocknet und im Vakuum eingeeengt. Umkristallisation aus  
20 Ethanol lieferten 3,03 g 17-Pseudoaglycon von Spinosyn A/D und Mutterlauge (C). Das Filtrat (B) wurde mehrmals mit Dichlormethan extrahiert. Die Extrakte wurden nacheinander mit gesätt. NaHCO<sub>3</sub>-Lösung und gesätt. NaCl-Lösung gewaschen, über Na<sub>2</sub>SO<sub>4</sub> getrocknet und im Vakuum eingeeengt. Der Rückstand wurde vereint mit der im Vakuum eingeeengten Mutterlauge (C) und durch Säulenchromatographie an  
25 650 cm<sup>3</sup> Kieselgel (Eluent: Cyclohexan/Essigsäureethylester 1:1, dann 100 % Essigsäureethylester) aufgetrennt. Man erhielt neben weiteren 1,76 g 17-Pseudoaglycon von Spinosyn A/D 0,78 g (16 %) Aglycon von Spinosyn A. Die Gesamtausbeute von 17-Pseudoaglycon von Spinosyn A/D betrug 4,79 g (69 %). – a) 17-Pseudoaglycon von Spinosyn A/D (82 % A, 18 % D): DC:  $R_f$  (SiO<sub>2</sub>, Essigsäureethylester) = 0,48.  
30 – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta$  = 6,78 (s, 13-H); 5,88 (d, 5-H von Spinosyn A); 5,80 (m, 6-H von Spinosyn A); 5,49 (m, 5-H von Spinosyn D); 4,86 (d, 1'-H); 4,70 (m, 21-H);

4,32 (m, 9-H) u. a. - LC/MS: Elektrospray Positiv; Peak bei RT 40,7 min:  $m/z = 609$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 641$  (10 %)  $[M+NH_4+CH_3OH]^+$  (Pseudoaglycon von Spinosyn A); Peak bei RT 41,4 min:  $m/z = 623$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 655$  (8 %)  $[M+NH_4+CH_3OH]^+$  (Pseudoaglycon von Spinosyn D). – b) Aglycon von Spinosyn A: DC:  $R_f$  (SiO<sub>2</sub>, Essigsäureethylester) = 0,29. – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta = 6,80$  (s, 13-H); 5,89 (d, 5-H); 5,80 (m, 6-H); 4,70 (m, 21-H); 4,44 (m, 9-H) u. a. - LC/MS: Elektrospray Positiv; Peak bei RT 36,8 min:  $m/z = 420$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 452$  (10 %)  $[M+NH_4+CH_3OH]^+$ .

#### 10 Darstellung des Aglycon von Spinosyn A/D

4,30 g (7,29 mmol) Pseudoaglycon von Spinosyn A/D wurden in 190 ml Methanol gelöst und mit 285 ml 7,2 N H<sub>2</sub>SO<sub>4</sub> versetzt. Nach 3 h Erwärmen unter Rückfluß wurde die abgekühlte Reaktionsmischung vorsichtig in 1700 ml gesättigte NaHCO<sub>3</sub>-Lösung gegeben. Man extrahierte mit Diethylether, wusch die Extrakte nacheinander mit gesätt. NaHCO<sub>3</sub>-Lösung und gesätt. NaCl-Lösung, trocknete über Na<sub>2</sub>SO<sub>4</sub> und engte im Vakuum ein. Nach Säulenchromatographie dieses Feststoffes an 650 cm<sup>3</sup> Kieselgel (Eluent: Cyclohexan/Essigsäureethylester 1:2, dann 100 % Essigsäureethylester) erhielt man 1,88 g (64 %) Aglycon von Spinosyn A. – DC:  $R_f$  (SiO<sub>2</sub>, Essigsäureethylester) = 0,29. – <sup>1</sup>H-NMR: CDCl<sub>3</sub>,  $\delta = 6,80$  (s, 13-H); 5,89 (d, 5-H); 5,80 (m, 6-H); 4,70 (m, 21-H); 4,44 (m, 9-H) u. a. – LC/MS: Elektrospray Positiv; Peak bei RT 36,6 min:  $m/z = 420$  (100 %)  $[M+NH_4]^+$ ,  $m/z = 452$  (14 %)  $[M+NH_4+CH_3OH]^+$  (Aglycon von Spinosyn A).

#### 25 Forosaminylierung des Spinosyn-Aglycons und Anknüpfung eines Trimethyl-Rhamnosezuckers an das Spinosyn-Aglycon durch Biotransformation mit einem rekombinanten *Streptomyces* Stamm, der Spinosyn-Zuckerbiosynthesegene heterolog exprimiert

30 20 ml R5A Medium (Illing et al., 1989) mit 5 µg Apramycin / ml wurde mit Mycel des rekombinanten Stammes *S. albus* (165-1) oder *S. albus* (165-8) angeimpft und

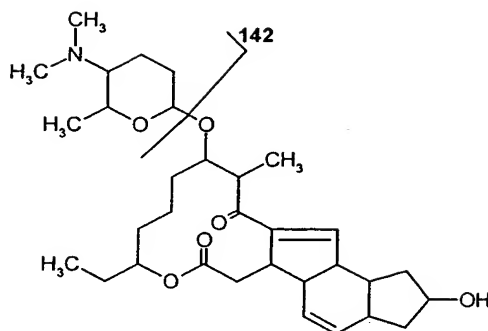


24 h aerob bei 28°C bebrütet. Diese Kultur wurde mit 50 µg/ml des hergestellten Spinosyn-Aglycons (100 µl einer 1 %igen Stammlösung in Methanol; Herstellung siehe Kapitel "Herstellung des Spinosyn-Aglycons und 17-Pseudoaglycons aus Tracer<sup>®</sup>", „Gewinnung von Spinosyn A/D aus Tracer<sup>®</sup>" und „Darstellung des Aglycon von Spinosyn A/D“) versetzt und ca. 120 h bei 28°C aerob inkubiert. Als Kontrolle wurde in gleicher Weise *S. albus* (pEBZ340; pOJ446-Vektor mit einem ca. 1,8 kb großen Spinosyn-PKS tragenden DNA-Fragment aus Cosmid 16-1-8) kultiviert und mit Spinosyn-Aglycon versetzt. Die Kulturen wurden nach Inkubation zur Abtrennung von Zellmycel zentrifugiert und der Überstand (20 ml) wurde mit 25 ml Methanol versetzt.

Je 35 ml des mit Methanol versetzten Kulturüberstandes wurden lyophilisiert, mit 15 ml Wasser aufgenommen und zweimal mit je 10 ml Essigsäureethylester extrahiert. Die vereinten organischen Phasen wurden zur Trockene eingengt und mit 350 µl Methanol aufgenommen. Ein Aliquot dieser Extrakte wurde mittels LC/MS mit Elektrospray Positiv-Ionisation untersucht.

Der Kulturüberstand der Anzucht von *S. albus* (165-1) enthielt eine Verbindung mit dem Molekulargewicht eines forosaminylierten Aglycons von Spinosyn A sowie Spinosyn A.

Peak 1: RT = 41,0 min: m/z = 544 (100 %) [M+H]<sup>+</sup>, m/z = 576 (20 %) [M+H+CH<sub>3</sub>OH]<sup>+</sup> (Forosaminyliertes Aglycon von Spinosyn A); LC/MS/MS: m/z = 142 (30 %) (Forosamin-Fragment).



Forosaminyliertes Aglycon von Spinosyn A  
 Molekulargewicht = 543  
 Summenformel =  $C_{32}H_{49}NO_6$

Peak 2: RT = 44,2 min:  $m/z = 733$  (100 %)  $[M+H]^+$  (Spinosyn A); LC/MS/MS:  $m/z = 142$  (21 %) (Forosamin-Fragment).

5

Der Kulturüberstand von *S. albus* (165-8) enthielt eine Verbindung mit dem Molekulargewicht eines Forosaminylierten Aglycons von Spinosyn A.

Peak 1: RT = 40,9 min:  $m/z = 544$  (100 %)  $[M+H]^+$ ,  $m/z = 576$  (20 %)  $[M+H+CH_3OH]^+$  (Forosaminyliertes Aglycon von Spinosyn A); LC/MS/MS:  $m/z = 142$  (39 %) (Forosamin-Fragment).

10

Der Kulturüberstand von *S. albus* (pEBZ340) enthielt keine Verbindungen mit MW 543 und kein Spinosyn A.

15

Forosaminylierung des Spinosyn-17-Pseudoaglycons durch Biotransformation mit einem rekombinanten *Streptomyces* Stamm, der Spinosyn Zuckerbiosynthesegene heterolog exprimiert

20

20 ml R5A Medium (Illing et al., 1989) mit 5  $\mu$ g Apramycin / ml wurde mit Mycel des rekombinanten Stammes *S. albus* (165-1) oder *S. albus* (165-8) angeimpft und 24 h aerob bei 28°C bebrütet. Diese Kultur wurde mit 50  $\mu$ g/ml des hergestellten 17-Pseudoaglycons von Spinosyn (100  $\mu$ l einer 1 %igen Stammlösung in Methanol; Herstellung siehe Kapitel "Herstellung des Spinosyn-Aglycons und 17-Pseudo-

aglycons aus Tracer<sup>®</sup>“, „Gewinnung von Spinosyn A/D aus Tracer<sup>®</sup>“ und „Darstellung des 17-Pseudoaglycon von Spinosyn A/D“) versetzt und ca. 120 h bei 28°C aerob inkubiert. Die Kulturen wurden nach Inkubation zur Abtrennung von Zellmycel zentrifugiert und der Überstand (20 ml) wurde mit 25 ml Methanol versetzt.

Je 35 ml des mit Methanol versetzten Kulturüberstandes wurden lyophilisiert, mit 15 ml Wasser aufgenommen und zweimal mit je 10 ml Essigsäureethylester extrahiert. Die vereinten organischen Phasen wurden zur Trockene eingengt und mit 350 µl Methanol aufgenommen. Ein Aliquot dieser Extrakte wurde mittels LC/MS mit Elektrospray Positiv-Ionisation untersucht.

Der Kulturüberstand von *S. albus* (165-1) enthielt Spuren von Spinosyn A und D.

Peak 1: RT = 44,2 min: m/z = 733 (100 %) [M+H]<sup>+</sup> (Spinosyn A); LC/MS/MS: m/z = 142 (8 %) (Forsamin-Fragment).

Peak 2: RT = 44,7 min: m/z = 747 (100 %) [M+H]<sup>+</sup> (Spinosyn D); LC/MS/MS: m/z = 142 (37 %) (Forsamin-Fragment).

Der Kulturüberstand von *S. albus* (165-8) enthielt Spuren von Spinosyn A und D.

Peak 1: RT = 44,1 min: m/z = 733 (100 %) [M+H]<sup>+</sup> (Spinosyn A).

Peak 2: RT = 44,7 min: m/z = 747 (100 %) [M+H]<sup>+</sup> (Spinosyn D).

#### Hinterlegung von Mikroorganismen

Folgende Mikroorganismen und Plasmide sind bei der Deutschen Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D- 38124 Braunschweig, in Übereinstimmung mit den Bestimmungen des Budapester Vertrages hinterlegt worden.

Mikroorganismus und Plasmid

Hinterlegungsnummer

	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-1-8	DSM 12961
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-2-2	DSM 12962
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-59-1	DSM 12963
5	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-59-6	DSM 12964
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 16-59-8	DSM 12965
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 165-1	DSM 13005
	<i>E. coli</i> XL1-Blue MRF' mit Cosmid 165-8	DSM 13007
	<i>E. coli</i> DH10B mit dem BAC Shuttle-Klon P8 / G11	DSM 13012
10	<i>E. coli</i> DH10B mit dem BAC Shuttle-Klon P11 / B10	DSM 13011
	<i>E. coli</i> DH10B mit dem BAC Shuttle-Klon P11 / G6	DSM 13010

Literatur

- 15 Biermann M., Logan R., O'Brien K., Seno E.T., Nagaraja R., Schoner B.E. (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene 116: 43-49
- 20 Chater K., Wilde L. (1980) *Streptomyces albus* G Mutants defective in the *SalGI* restriction-modification system. J. Gen. Microbiol. 116: 323-334.
- Devereux J., Haeberli P., Smithies (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Research 12: 387-395.
- 25 Mertz F., Yao R.C. (1990) *Saccharopolyspora spinosa* sp. nov. isolated from soil collected in a sugar mill rum still. Int. J. Syst. Bacteriol. 40: 34-39.
- 30 Creemer L.C., Kirst H.A., Paschal J.W. (1998) Conversion of spinosyn A and spinosyn D to their respective 9- and 17-pseudoaglycones and their aglycones. J. Antibiotics 51: 795-800.

Fishman S.E., Hershberger C.L. (1983) Amplified DNA in *Streptomyces fradiae*. J. Bacteriol. 155: 459-466.

5 Hopwood D.A., Bibb M.J., Chater K.F., Kieser T., Bruton C.J., Kieser H.M., Lydiate D.J., Smith C.P., Ward J.M., Schrempf H. (1985) Genetic manipulation of *streptomyces* a laboratory manual, The John Innes Foundation, Norwich, 1985.

Illing G.T., Normansell I.D., Peberdy J.F. (1989) Protoplast isolation and regeneration in *Streptomyces clavuligerus*. J. Gen. Microbiol. 135: 2289-2297.

10

Kim U.J. Birren B.W. Slepak T. Mancino: V. Boysen C. Kang H.L. Simon M.I. Shizuya H. (1996), Construction and characterization of a human bacterial artificial chromosome library. Genomics 34: 213-218.

15

Kirst H.A., Michel K.H., Martin J.W., Creemer L.C., Chao E.H., Yao R.C., Nakatsukasa W.M., Boeck L.D., Occolowitzh J.L., Paschal J.W., Deeter J.B., Jones N.D., Thompson G.D. (1991) Tetrahedron Lett. 32: 4839-4842.

20

Motamedi H., Shafiee A., Sheng-Jian C. (1995) Integrative vectors for heterologous gene expression in *Streptomyces* spp. Gene 160: 25-31.

25

Muyrers J.P.P., Zhang Y., Testa G., Stewart A.F. (1999) Rapid modification of bacterial artificial chromosomes by ET-recombination. Nucleic Acids Research 27: 1555-1557.

Nakatsukasa W.M., Mabe J.A., Michel K.H., Martin J.W., Paschal J.W., Elzey T.K. (1990) Abstracts of 2<sup>nd</sup> Int. Conf. on Biotechnology of Microbial Prod.: P-21.

30

Sambrook J., Fritsch E.F., Maniatis T. (1989) Molecular cloning – A laboratory manual. Second edition. Cold Spring Harbor Laboratory Press.

Shizuya H., Birren B., Kim U.-J., Mancino V., Slepak T., Tachiiri Y., Simon M.I. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc. Natl. Acad. Sci. USA 89: 8794-8797.

5

Smokvina T., Mazodier P., Boccard F., Thompson C.J., Guérineau M. (1990) Construction of a series of pSAM2-based integrative vectors for use in actinomycetes. Gene 94: 53-59.

**Patentansprüche**

1. Nucleinsäure, welche zumindest eine Region umfasst, die für eine Enzymaktivität codiert, welche an der Biosynthese von Spinosynen beteiligt ist.
2. Nucleinsäure gemäß Anspruch 1, dadurch gekennzeichnet, dass es sich um einzelsträngige oder doppelsträngige DNA oder RNA handelt.
3. Nucleinsäure gemäß Anspruch 2, dadurch gekennzeichnet, dass es sich um ein DNA-Fragment handelt.
4. Nucleinsäure gemäß Anspruch 3, dadurch gekennzeichnet, dass sie alle Regionen umfasst, die für Enzymaktivitäten codieren, welche an der Biosynthese von Spinosynen beteiligt sind.
5. Nucleinsäure gemäß einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, dass es sich um Enzymaktivitäten von Polyketidsynthasen, Methyltransferasen, Glycosyltransferasen, Epimerasen, Aminotransferasen, Dimethyltransferasen, Reduktasen, Dehydratasen und/oder Cyclisierungsenzymen handelt.
6. Nucleinsäure gemäß einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass sie aus *Saccharopolyspora spinosa* stammt.
7. Nucleinsäure gemäß Anspruch 1, umfassend zumindest eine Sequenz ausgewählt aus
  - (a) den Sequenzen gemäß SEQ ID NOS: 1, 2, 3, 4, 5, 6, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 52 oder 54,

- (b) zumindest 14 Basenpaare langen Teilsequenzen der unter (a) definierten Sequenzen,
- (c) Sequenzen, welche an die unter (a) definierten Sequenzen hybridisieren
- (d) Sequenzen, welche eine zumindest 70 %ige Identität zu den unter (a) definierten Sequenzen aufweisen,
- (e) Sequenzen, welche zu den unter (a) definierten Sequenzen komplementär sind, und
- (g) Sequenzen, welche aufgrund der Degeneriertheit des genetischen Codes für dieselbe Aminosäuresequenz kodieren wie die unter (a) bis (d) definierten Sequenzen.
8. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie die Sequenz gemäß SEQ ID NOS: 1 bis 6 umfasst.
9. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie die Sequenz gemäß SEQ ID NO: 4 umfasst.
10. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie die Sequenz gemäß SEQ ID NOS: 5 und 6 umfasst.
11. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie zumindest eine Sequenz gemäß SEQ ID NOS: 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37 oder 39 umfasst.
12. Nucleinsäure gemäß Anspruch 7, dadurch gekennzeichnet, dass sie zumindest eine Sequenz gemäß SEQ ID NOS: 41, 43, 45, 47 oder 49 umfasst.



13. Regulatorische Region, welche die Transkription einer Nukleinsäure gemäß einem der Ansprüche 1 bis 7 in *Saccharopolyspora spinosa* kontrolliert.
- 5 14. DNA-Konstrukt umfassend eine Nucleinsäure gemäß einem der Ansprüche 1 bis 12 und zumindest einen heterologen Promotor.
- 10 15. Vektor umfassend zumindest eine Nucleinsäure gemäß einem der Ansprüche 1 bis 12, eine regulatorische Region gemäß Anspruch 13 oder ein DNA-Konstrukt gemäß Anspruch 14.
- 15 16. Vektor gemäß Anspruch 15, dadurch gekennzeichnet, dass die Nucleinsäure funktionell mit regulatorischen Sequenzen verknüpft ist, welche die Expression der codierenden Regionen der Nucleinsäure in pro- oder eukaryotischen Zellen gewährleisten.
- 20 17. Vektor gemäß Anspruch 15 oder 16, dadurch gekennzeichnet, dass es sich um einen BAC-Vektor, PAC-Vektor oder einen zu BAC- oder PAC-Vektoren funktionell gleichwertigen Vektor handelt.
- 25 18. Vektor gemäß Anspruch 17, dadurch gekennzeichnet, dass es sich um einen Vektor handelt, der den BAC-Klonen mit den Hinterlegungsnummern DSM 13010, DSM 13011 oder DSM 13012 entspricht.
- 30 19. Vektor gemäß einem der Ansprüche 15 bis 18, dadurch gekennzeichnet, dass es sich um einen Shuttle-Vektor handelt, der sowohl in Prokaryonten als auch in Eukaryonten übertragbar ist.
20. Vektor gemäß einem der Ansprüche 15 bis 19, dadurch gekennzeichnet, dass es sich um einen Shuttle-Vektor handelt, der sowohl in Gram-negative und in Gram-positive Bakterien als auch in Archea übertragbar ist.

21. Vektor gemäß einem der Ansprüche 15 bis 19, dadurch gekennzeichnet, dass es sich um einen Shuttle-Vektor handelt, der sowohl in *Escherichia coli* als auch in Actinomyceten übertragbar ist.
- 5
22. Vektor gemäß Anspruch 21, dadurch gekennzeichnet, dass es sich um einen Shuttle-Vektor handelt, der sowohl in *Escherichia coli* als auch in *Streptomyces* übertragbar ist.
- 10
23. Vektor gemäß einem der Ansprüche 15 bis 22, dadurch gekennzeichnet, dass er in einem Prokaryonten autonom replizierbar ist.
24. Vektor gemäß einem der Ansprüche 15 bis 22, dadurch gekennzeichnet, dass er in einem Prokaryonten unter Beteiligung des Phagen  $\Phi$ C31-, des pSAM2- oder des Mini-Circle-Integrationsmechanismus ins Genom integrierbar ist.
- 15
25. Vektor gemäß einem der Ansprüche 15 bis 22, dadurch gekennzeichnet, dass er in einem Prokaryonten durch RecA-vermittelte Rekombination ins Genom integrierbar ist.
- 20
26. Vektor gemäß einem der Ansprüche 15 bis 22, dadurch gekennzeichnet, dass er in einem Prokaryonten durch RecE- und RecT-vermittelte Rekombination ins Genom integrierbar ist.
- 25
27. Wirtszelle enthaltend eine Nucleinsäure gemäß einem der Ansprüche 1 bis 12, eine regulatorische Region gemäß Anspruch 13, ein DNA-Konstrukt gemäß Anspruch 14 oder zumindest einen Vektor gemäß einem der Ansprüche 15 bis 26.
- 30
28. Wirtszelle gemäß Anspruch 27, dadurch gekennzeichnet, dass es sich um eine pro- oder eukaryotische Zelle handelt.

29. Wirtszelle gemäß Anspruch 28, dadurch gekennzeichnet, dass die prokaryotische Zelle zur Gruppe der Actinomyceten, bevorzugt zur Gruppe der Streptomyceten gehört.
- 5 30. Wirtszelle gemäß Anspruch 28, dadurch gekennzeichnet, dass die eukaryotische Zelle eine Pflanzenzelle ist.
- 10 31. Polypeptid, welches von einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 codiert wird.
32. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer Methyltransferase aufweist.
- 15 33. Polypeptid gemäß Anspruch 32, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NOS: 8, 12, 14, 18 oder 20, oder eine Teilsequenz davon aufweist.
- 20 34. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer Glycosyltransferase aufweist.
35. Polypeptid gemäß Anspruch 34, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NOS: 10 oder 30, oder eine Teilsequenz davon aufweist.
- 25 36. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität eines C-C verknüpfenden Enzyms, das Cyclisierungsreaktionen durchführt, aufweist.
- 30 37. Polypeptid gemäß Anspruch 36, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 16 oder eine Teilsequenz davon aufweist.

38. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität eines Enzyms, das an Cyclisierungsreaktionen beteiligt ist, aufweist.
- 5 39. Polypeptid gemäß Anspruch 38, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 22 oder eine Teilsequenz davon aufweist.
40. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer 2,3-Reduktase aufweist.
- 10 41. Polypeptid gemäß Anspruch 40, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 24 oder eine Teilsequenz davon aufweist.
42. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer 2,3-Dehydratase aufweist.
- 15 43. Polypeptid gemäß Anspruch 42, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 26 oder eine Teilsequenz davon aufweist.
- 20 44. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer Thioesterase aufweist.
45. Polypeptid gemäß Anspruch 44, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 28 oder eine Teilsequenz davon aufweist.
- 25 46. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer 3,4-Dehydratase aufweist.
- 30 47. Polypeptid gemäß Anspruch 46, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 32 oder eine Teilsequenz davon aufweist.

48. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer 4-Aminotransferase aufweist.
- 5 49. Polypeptid gemäß Anspruch 48, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 34 oder eine Teilsequenz davon aufweist.
50. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer N-Dimethyltransferase aufweist.
- 10 51. Polypeptid gemäß Anspruch 50, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 36 oder eine Teilsequenz davon aufweist.
52. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer 3,4-Reduktase aufweist.
- 15 53. Polypeptid gemäß Anspruch 52, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 38 oder eine Teilsequenz davon aufweist.
- 20 54. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität eines Transkriptions-Regulators aufweist.
55. Polypeptid gemäß Anspruch 54, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 40 oder eine Teilsequenz davon aufweist.
- 25 56. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer Polyketidsynthase aufweist.
- 30 57. Polypeptid gemäß Anspruch 56, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NOS: 42, 44, 46, 48 oder 50, oder eine Teilsequenz davon aufweist.

58. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer Glucose-Dehydratase aufweist.
59. Polypeptid gemäß Anspruch 58, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 53 aufweist.
60. Polypeptid gemäß Anspruch 31, dadurch gekennzeichnet, dass es die Aktivität einer 3,5-Epimerase aufweist.
61. Polypeptid gemäß Anspruch 60, dadurch gekennzeichnet, dass es die Aminosäuresequenz gemäß SEQ ID NO: 55 aufweist.
62. Enzyme, die an Cyclisierungsreaktionen beteiligt sind, dadurch gekennzeichnet, dass sie die Aminosäuresequenz gemäß SEQ ID NO: 15 oder 22, oder eine Teilsequenz davon, welche zumindest noch eine Teilreaktion durchführen kann, umfassen oder eine mindestens 50 %ige Identität dazu auf Aminosäureebene aufweisen.
63. Antikörper, welcher spezifisch mit einem Polypeptid gemäß einem der Ansprüche 31 bis 62 reagiert.
64. Verfahren zur Herstellung einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7, umfassend die folgenden Schritte:
- (a) Vollständige chemische Synthese auf an sich bekannte Weise oder
  - (b) chemische Synthese von Oligonucleotiden, Markieren der Oligonucleotide, Hybridisieren der Oligonucleotide an DNA einer genomischen oder cDNA-Bank, die ausgehend von genomischer DNA bzw. mRNA aus *S. spinosa* hergestellt wurde, Selektieren von

positiven Klonen und Isolieren der hybridisierenden DNA aus positiven Klonen oder

(c) chemische Synthese von Oligonucleotiden und Amplifizierung der Ziel-DNA mittels PCR.

65. Verfahren zur Herstellung eines Polypeptids gemäß einem der Ansprüche 31 bis 62, umfassend die folgenden Schritte:

(a) Kultivieren einer Wirtszelle gemäß einem der Ansprüche 27 bis 30 unter Bedingungen, die die Expression der Nucleinsäure gemäß einem der Ansprüche 1 bis 7 gewährleisten, oder

(a1) Exprimieren einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 in einem *in vitro*-System, und

(b) Gewinnen des Polypeptids aus der Zelle, dem Kulturmedium oder dem *in vitro*-System.

66. Verfahren zum Herstellen von Spinosyn, Spinosyn-Vorstufen oder Spinosyn-Derivaten umfassend die folgenden Schritte:

(a) Kultivieren einer Wirtszelle gemäß einem der Ansprüche 27 bis 30 unter Bedingungen, die die Expression der Nucleinsäure gemäß einem der Ansprüche 1 bis 7 gewährleisten und

(b) Gewinnen des Spinosyns, der Spinosyn-Vorstufe oder des Spinosyn-Derivates aus der Zelle oder dem Kulturmedium.

67. Verfahren zum Herstellen von Spinosyn-Derivaten, einschließlich Spinosyn-Vorstufen, umfassend die folgenden Schritte:

- 5
- (a) Austauschen zumindest einer Modul-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Modul-codierende Nucleinsäuresequenz gemäß Anspruch 7, oder
- 10
- (b) Austauschen zumindest einer Modul-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Modul-codierende Nucleinsäuresequenz aus *S. spinosa*, oder
- 15
- (c) Austauschen zumindest einer Modul-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Modul-codierende Nucleinsäuresequenz aus einem anderen Organismus als *S. spinosa*, oder
- 20
- (d) Austauschen zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Domänen-codierende Nucleinsäuresequenz gemäß Anspruch 7, oder
- 25
- (e) Austauschen zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Domänen-codierende Nucleinsäuresequenz aus *S. spinosa*, oder
- 30
- (f) Austauschen zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen zumindest eine andere Domänen-codierende Nucleinsäuresequenz aus einem anderen Organismus als *S. spinosa*, oder
- (g) Austauschen einer ersten Acyltransferase-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen eine zweite Acyltransferase-codierende Nucleinsäuresequenz gemäß Anspruch 7, wobei die zweite



Acyltransferase eine unterschiedliche Substratspezifität aufweist als die erste Acyltransferase, oder

5

- (h) Austauschen einer ersten Acyltransferase-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen eine zweite Acyltransferase-codierende Nucleinsäuresequenz aus *S. spinosa*, wobei die zweite Acyltransferase eine unterschiedliche Substratspezifität aufweist als die erste Acyltransferase, oder

10

- (i) Austauschen einer ersten Acyltransferase-codierenden Nucleinsäuresequenz gemäß Anspruch 7 gegen eine zweite Acyltransferase-codierende Nucleinsäuresequenz aus einem anderen Organismus als *S. spinosa*, wobei die zweite Acyltransferase eine unterschiedliche Substratspezifität aufweist als die erste Acyltransferase, oder

15

- (j) Deletieren zumindest einer Domänen-codierender Nucleinsäuresequenz gemäß Anspruch 7, oder

20

- (k) Integrieren zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7 in eine Modul-codierende Nucleinsäuresequenz gemäß Anspruch 7, oder

25

- (l) Mutagenisieren zumindest einer Domänen-codierenden Nucleinsäuresequenz gemäß Anspruch 7,

und Exprimieren der rekombinierten Nucleinsäuresequenz in einer Wirtszelle unter Bedingungen, welche die Synthese eines Spinosyn-Derivates oder einer Spinosyn-Vorstufe erlauben.

68. Verwendung einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 zum Identifizieren, Inaktivieren und/oder Modifizieren von Genen der Spinosyn-Biosynthese.
- 5 69. Verwendung einer Nucleinsäure gemäß einem der Ansprüche 1 bis 7 zum Erzeugen einer Bibliothek aus Polyketidsynthasen.
70. Verfahren zum Anfügen eines Forosamin-Zuckerrestes an das Spinosyn-Aglycon oder an das Spinosyn-17-Pseudoaglycon oder an ein Polyketid-Aglycon, umfassend die folgenden Schritte:
- 10
- (a) Übertragen einer Nucleinsäure gemäß SEQ ID NOS: 23, 25, 29, 31, 33, 35 und 37 in eine Wirtszelle, die das Spinosyn-Aglycon oder das Spinosyn-17-Pseudoaglycon oder das Polyketid-Aglycon herstellen kann, oder
- 15
- (a1) Übertragen einer Nucleinsäure gemäß SEQ ID NOS: 23, 25, 29, 31, 33, 35 und 37 in eine Wirtszelle, die das Spinosyn-Aglycon oder das Spinosyn-17-Pseudoaglycon oder das Polyketid-Aglycon nicht herstellen kann und Zufügen des Spinosyn-Aglycons oder des Spinosyn-17-Pseudoaglycons oder des Polyketid-Aglycons zum Kulturmedium, und
- 20
- (b) Kultivieren der Wirtszelle unter Bedingungen, die zu einem aktiven Stoffwechsel der Zelle führen.
- 25
71. Verfahren zum Anfügen eines Trimethyl-Rhamnose-Zuckerrestes an das Spinosyn-Aglycon oder an das Spinosyn-17-Pseudoaglycon oder an ein Polyketid-Aglycon, umfassend die folgenden Schritte:
- 30

- (a) Übertragen einer Nucleinsäure gemäß SEQ ID NO: 7, 9, 11, 13, 17 und/oder 19 in eine Wirtszelle, die das Spinosyn-Aglycon oder das Spinosyn-17-Pseudoaglycon oder das Polyketid-Aglycon herstellen kann, oder

5

- (a1) Übertragen einer Nucleinsäure gemäß SEQ ID NO: 7, 9, 11, 13, 17 und/oder 19 in eine Wirtszelle, die das Spinosyn-Aglycon oder das Spinosyn-17-Pseudoaglycon oder das Polyketid-Aglycon nicht herstellen kann und Zufügen des Spinosyn-Aglycons oder des Spinosyn-17-Pseudoaglycons oder des Polyketid-Aglycons zum Kulturmedium, und

10

- (b) Kultivieren der Wirtszelle unter Bedingungen, die zu einem aktiven Stoffwechsel der Zelle führen.

15

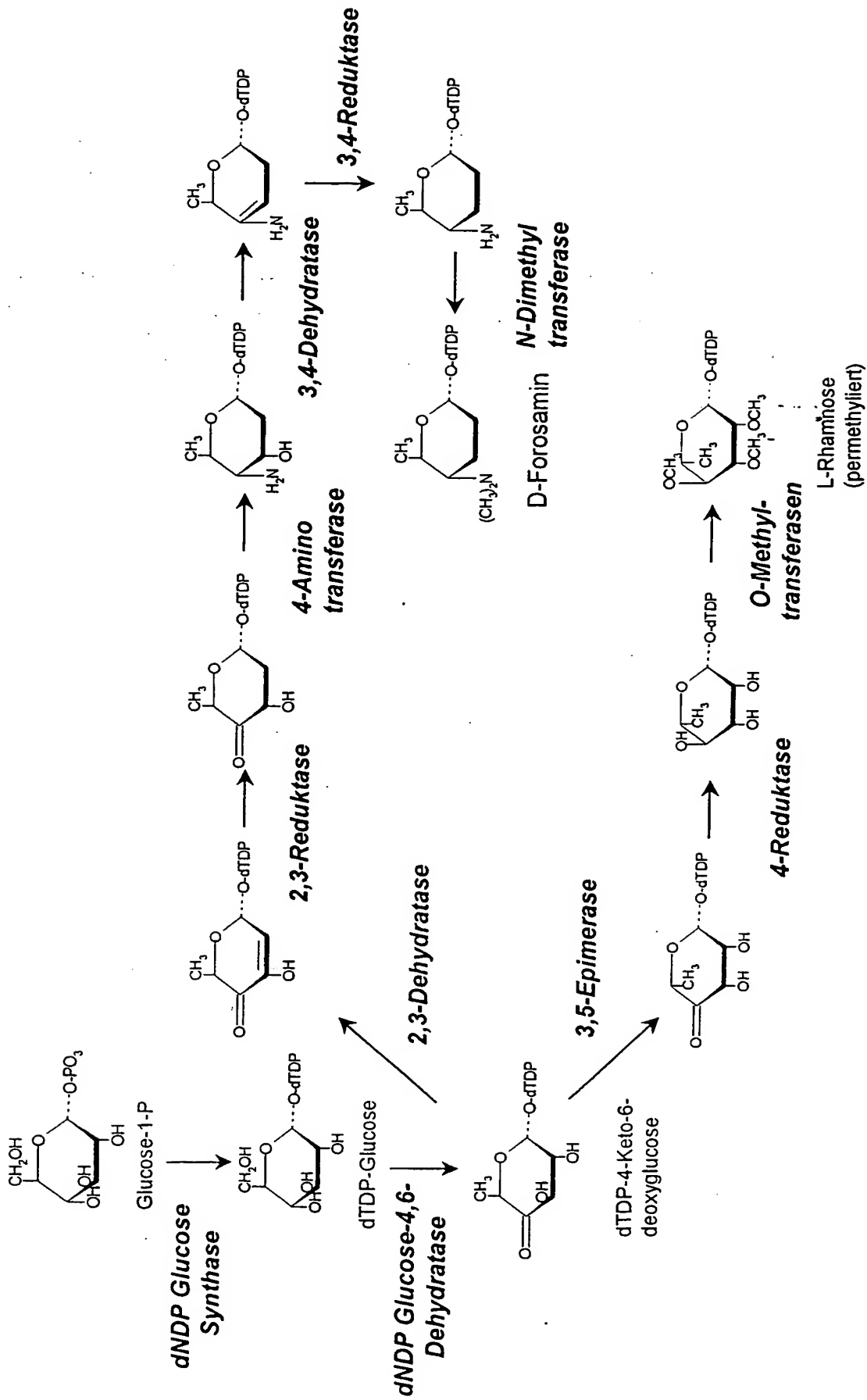
72. Verfahren gemäß Anspruch 71, dadurch gekennzeichnet, dass im Schritt (a) Nucleinsäuren gemäß SEQ ID NOS: 9, 11, 13 und 17 übertragen werden.

Nucleinsäuren, die für Enzymaktivitäten der Spinosyn-Biosynthese codieren

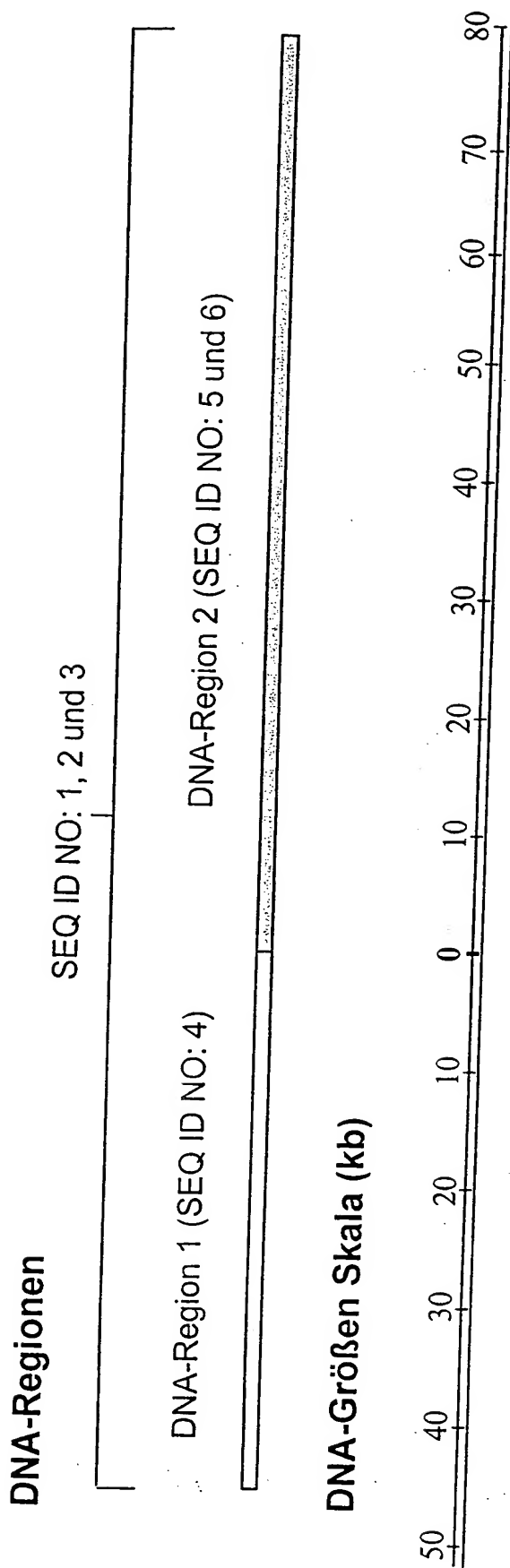
Z u s a m m e n f a s s u n g

Die vorliegende Erfindung betrifft Nucleinsäuren, die für Enzymaktivitäten der Spinosyn-Biosynthese codieren, sowie die entsprechenden Enzyme per se. Weiterhin betrifft die Erfindung Verfahren zum Herstellen von Spinosyn-Derivaten und -Vorstufen.

**Abbildung 1: Metabolismus für die Biosynthese der Nucleosyn Zucker**  
**D-Forosamin und 2, 3, 4 Tri-O-Methyl-L-Rhamnose**



**Abbildung 2:** Lage, der **Spinosyn Biosynthese** oder indirekt beteiligten DNA-Regionen 1 und 2 sowie Lage und Bezeichnung der zur Sequenzierung herangezogenen überlappenden Cosmidklone



**Cosmid-Bezeichnung, sowie Größe, Position und Überlappung von Insert-DNA**

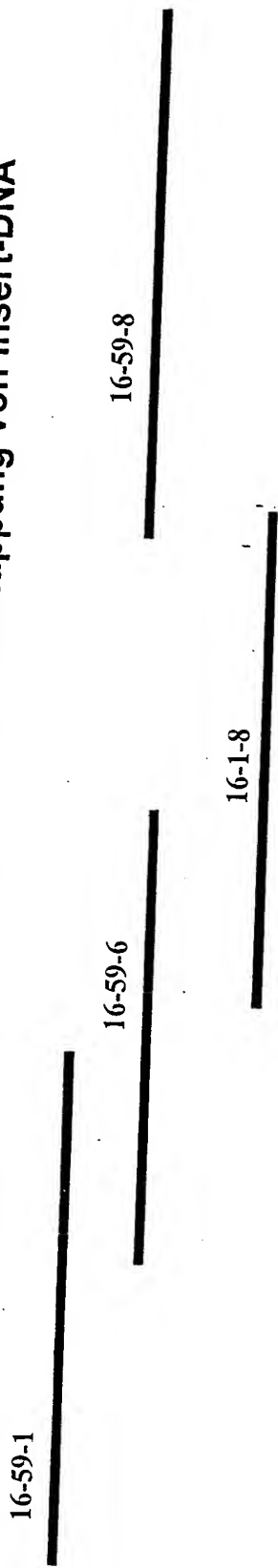
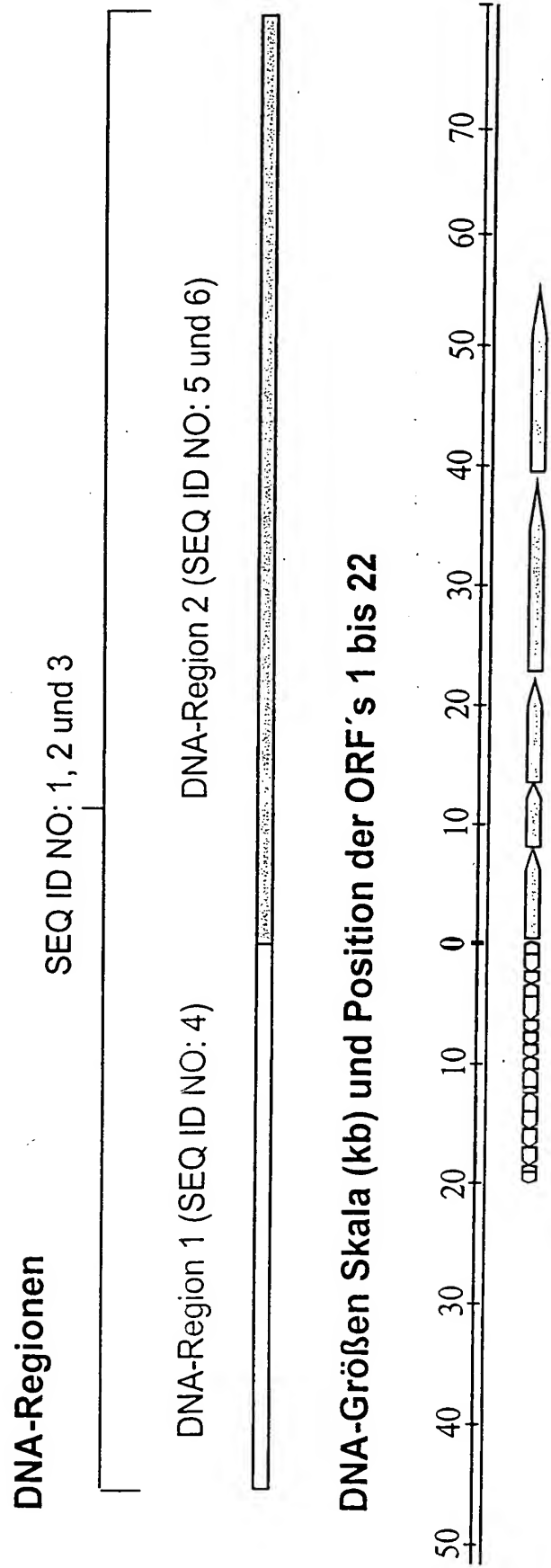


Abbildung 3: Z...otransformation herange...ne Cosmidklone



Cosmid-Bezeichnung, sowie Größe und Position von Insert-DNA

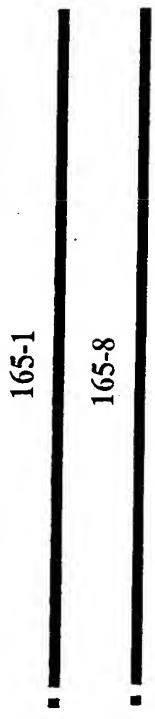
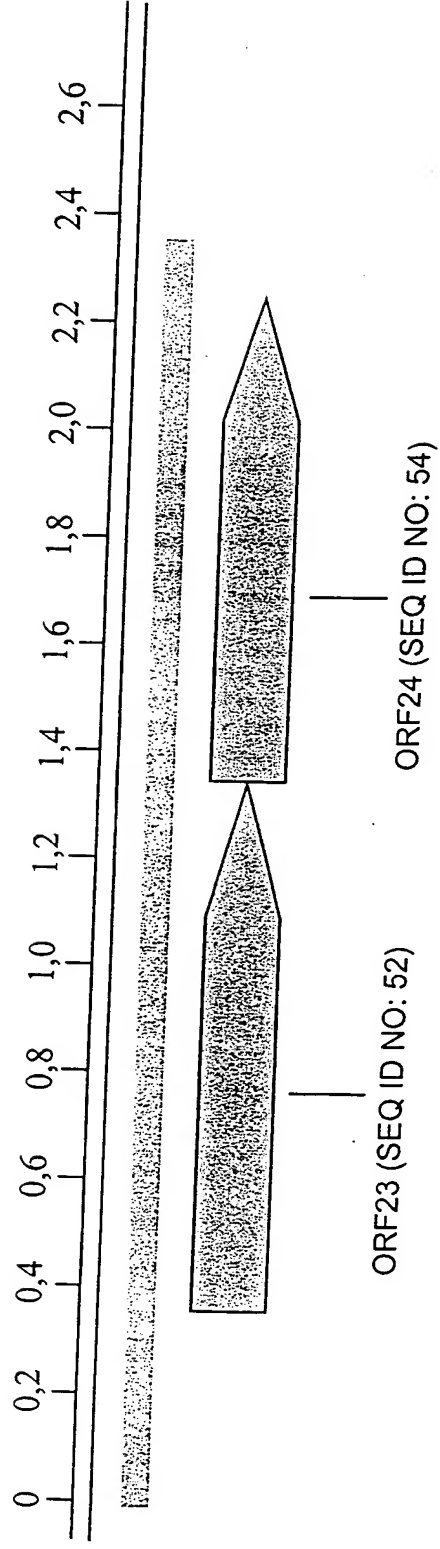


Abbildung 4: Darstellung offener Leseraster  
der DNA-Region 3 (SEQ ID NO: 51) auf Cosmid 16-2-2

SEQ ID NO: 51

DNA-Größen Skala (kb) und Nucleotidposition





**Abbildung 5:** Darstellung offener Leserahmen der DNA-Regionen 1 und 2, die direkt oder indirekt an der Spinosyn-Biosynthese involviert sind

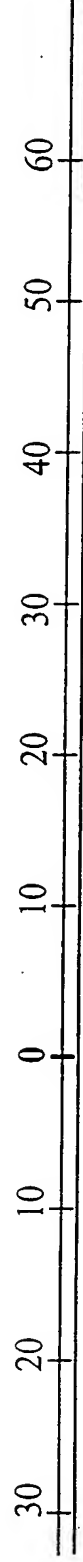
## DNA-Region

DNA-Region 1 (Ausschnitt)

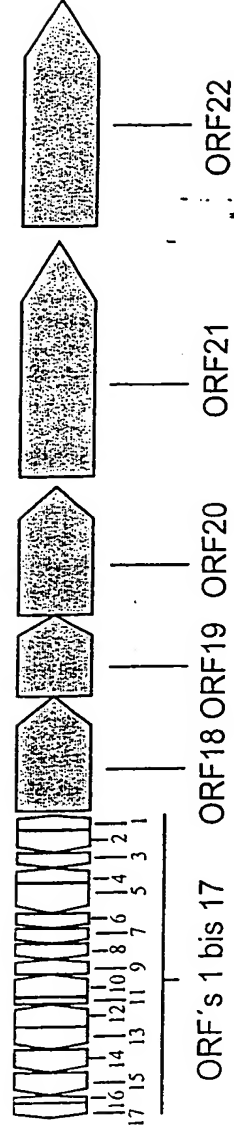
DNA-Region 2 (Ausschnitt)



## DNA-Größen Skala (kb) und Nucleotidposition

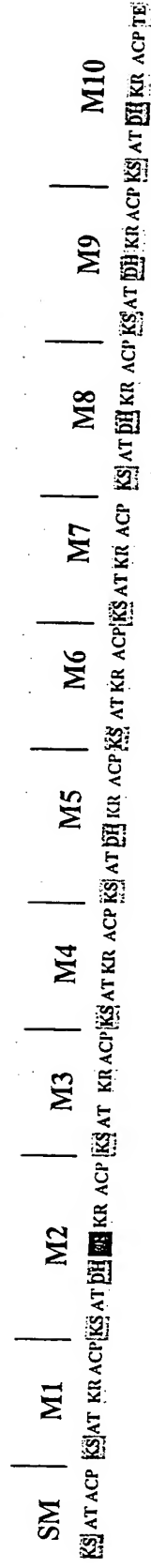


## Offene Leserahmen (ORF's)

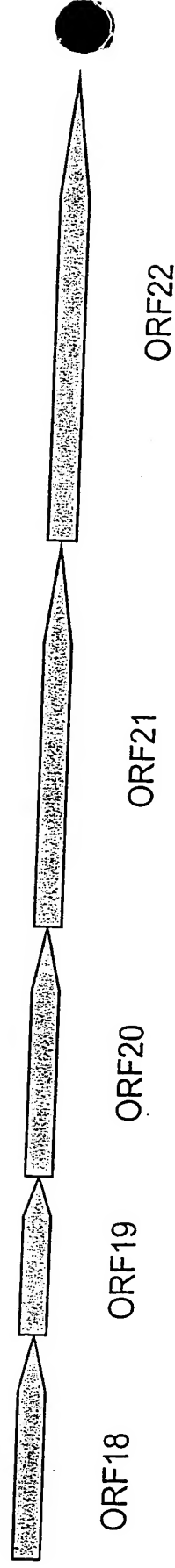


**Abbildung 6: Darstellung offener Leserahmen und DNA-Region 2 und ableitbare Polyketidsynthase-Module und -Domänen**

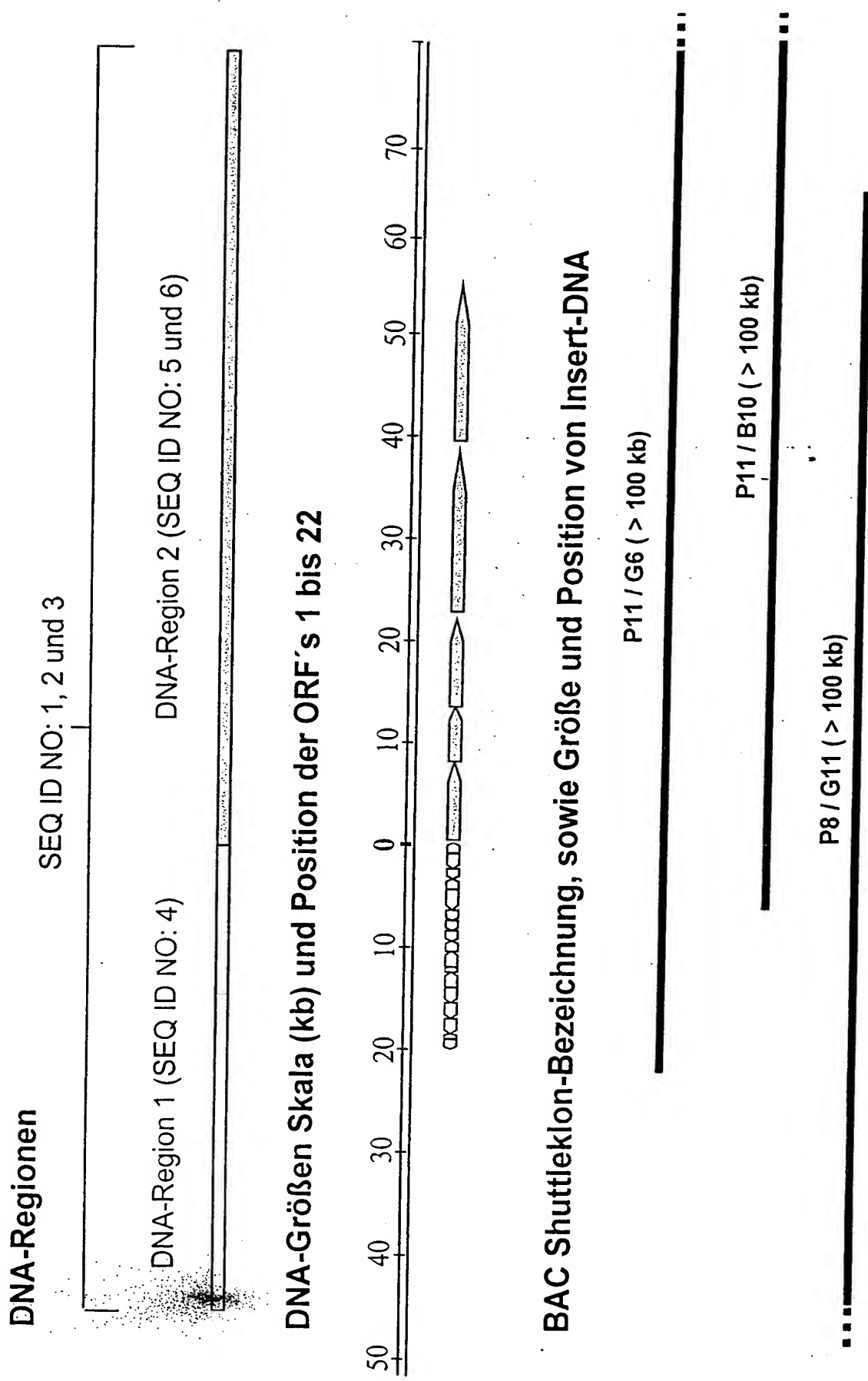
### Module und Domänen



### Offene Leserahmen



**Abbildung 7: Darstellung von Spinosyn Biosynthesegenen in BAC Shuttleklonen**



# SEQUENZPROTOKOLL

<110> Bayer Aktiengesellschaft

<120> Nucleinsäuren, die für Enzymaktivitäten der  
Spinosyn-Biosynthese codieren

<130> Le A 33 955

<140>

<141>

<160> 55

<170> PatentIn Ver. 2.1

<210> 1

<211> 50000

<212> DNA

<213> Saccharopolyspora spinosa

<400> 1

```

ggcaccaccc cggggccccc gcctcactcg gagtcgatcc gcattccggt cacgaacatg 60
tgacccgctcg agtgcgatg cattagctcg gtcttcgect cgatcgccgc ctgctgcggg 120
gtcacgcgcg aggccagaa caccggcacg tcgcccggct ccatacgcac tgcgtcaccc 180
cagtcagggt catcaacgtt gtcgatgccc aacttctcag gggagccgat gtgcacgggg 240
cccccgtgca cctcgggatg cgcgatgctg atggtgacag cctccgccac ccggtccggg 300
ggcatgggac gcattgaaac cactagcttc ccccggaact ggctgcgggg cacacactca 360
cggttgctga ggtaaaccgg gccgttgatc ccctggtcga tgtgccggat cgggagtcgg 420
gcatcgagca gcaagtgtc gaaggagaag ctgcagccaa gcaggaacgc gaccatgtcg 480
tcggtccaca ggtcaacgac gtcactgcgc tctccacca gcgccccttc ccggtacacg 540
ctgtatttgg gcaggtcggt goggaggtca gcaccagggg cggtcagctt cggctccgga 600
tctcccgcgt cggtgacctc gaggatcggg caccgcttgg ggttcgctg ggccaaccgc 660
atgaagtcca acgcgtcatc ctggggcagc atcacagagt tggcctgcac gaatccatga 720
cagatcccgg acgtgggect cgcccacgca ttcgcgcggt tggcacgcca tgcctcgatg 780
ggctccgctg gctgcagatc aacgaactcc atcgatgact tccccttccc ggcgatgatg 840
acgaaggctg gcaaccgagg gttgcccgct acgggtggcc gtcactggtc tgccacccag 900
gggtcgccga gcaggctctg gtgaacgatt ccggagatca gctcggcatg cgcggcgacg 960
aaacgcagtg ggatcccgtc gtcgccttgg aacgcgggat cgtcaacgcc gccgtgcgt 1020
agacctgcgc gcagccccag ctcggtccac tcgcgcacca gatccccggt gactgccagg 1080
tttgcgctc gagccaggat ggccaccgca ccggtgatga cgtagcagcc caaattcgac 1140
acggatgcag gcagcagatg atccaccgtc gtggacgcga agaattcgcc ctcaacgctg 1200
acgcctcgcg gagtcagcgc ctcgactcg gcggcaatgt tgccgaacct gatctcgttg 1260
ccgttgtctc ccacggccag cgtcagagca ccgttgtccg tggcggcgcg aaggtagtcg 1320
tcggcaaaca gatcaccggc ggtgaccggc gtgccccaga tcagatgacg gccgcaaca 1380
gagttctgcc ccagcttctc gatggcgggc acgacgcgca actgcgcgcg gaactcccgg 1440
gcctcggccg cgttgagaaa tcgggactcc atcaaggtca cgcctccag gccgcagcg 1500
tcgaccagct tctccaccgg gccgacggcc tctgggtcga tgacgatggt cacctcgcta 1560

```

ccaagaccgg cgagcgtgcg cgcgagtgcc agagagccga tggggccgctc cacctcccc 1620  
 tgcgggaacc ggccttcctc gaccagaccc gtgacgatca acgccttttc accctgtcga 1680  
 cgcagcagag agcgcgcgat cgcaaggctg atcggctccg accgctcacc ccgcgccagc 1740  
 gtgtagaacg ggctcacgat gccggccggg ctgctcggcg gacgcattctc gatcgtggcc 1800  
 agtcgatcca ggctctcgaa ggccgggtgc gactcgggct tcatcagctc tccctcggta 1860  
 tacgacatat cccaaggact tcatccgcta cggagtcccg cctcggggcg ttgcttacgc 1920  
 ggtgaggacc atataccgta ttccactgtc caaggaaggg ggacgatgga accgaatcag 1980  
 ccacgcactg cgggcacgct gaacatgggt gagctgcacc atcgcccagc cgccgcagac 2040  
 tccgcgggtg ggctccggcaa cggatgctc gcacccctcg acgcacgcgt cgacaatctg 2100  
 acagatcggg tcttcgaggc catccgcgac agcatcgtaa gtcaggcact cccacctggg 2160  
 cgtcgtgtga gcgaggccaa gctcgcggcc gagctgaatg tcagcaagac ccccgtcgc 2220  
 gaggcgcttc tccgcttgcg gcacgtgggc ctggctcagg ctgccccgaa ggggctgcgc 2280  
 gtcacacgcg cctcgggtccg cgccattcgc gacgcctacg agcaccgggc ggggctggag 2340  
 agctcggcag cctggttcac cgcacatcgg gctaccagcg cccaacggga gaggtccac 2400  
 gaggtcgcca cctcatccct gcgcgcggca gagtcggcg actccgcca gttccgcgc 2460  
 gaggatcgcg tcttccacca catcatcgcc gagttctgcg ggaacgagat cctccagcgc 2520  
 gcggtcgaca acgccttggt gctgaccgct gtactccgag aacgcgacac tccctcgacc 2580  
 ggtgactcca tctcctgcgc ctggagcac atcgcgacgg catccgccgt tctggctggc 2640  
 gacgccgacc gggcagcccg gaacaacgcc gaccacatcc tccacgtgat gtcgcttgte 2700  
 ctctcctcac acacgcagcg cgggaacgtg cccccgggaa cgagttaaac caactcccc 2760  
 accccctacc tcaccattcg accgaaggac gggctttcgt gaccgagcac ctggcgacca 2820  
 tcaccgctgc ggaaggagag atccccgttgc tegttagcga ccgcggcgctc gcagcaatcc 2880  
 gggagataga ccccgagctg gcggtcacga ccatcatgga cgtgatcgca ggagaccacc 2940  
 tcgcgcggct gcgtgagcgt gtcaagaacg cccccgacct cgccttcaca cgcgtggacg 3000  
 aggtgcggtt caccgcgccg taccgatccc cccggaagat ctggggcatc ggtctgaact 3060  
 acggcgatca cgcagcggac ctgagcaggt cggcgcccga ccagccggcg tcttcgtga 3120  
 agtgcgacca cagcatcatc ggcccggatg agccgatcgt gattccggcg cagagtgagc 3180  
 gcaccacgtc cgaggctgag atcgggctca tcatcggcag gacttgcgag aacgtcagcg 3240  
 agtccgaggc gatggactac gtgtgggggg tcacgaccat cctggaccag accgccgagg 3300  
 acatcttgcg gatcaatccg cgctatctca cccgctcgaa gaacttcggg acgttcttct 3360  
 gcttcggacc ggagctcgta ccgctcgatg aggtgataga gaagttcgaa agcttgagc 3420  
 acatcgagat ctcgacggtc aagaacggcg atgagttcag gaccaacacg gtcgcccaca 3480  
 tgaccacaa gcccgccagc ctgctgagct tccactcgaa gatgatgccg ctgtttccgg 3540  
 gagacatcat ctccaccgga accctggcg cgctcgtcat cggcgacgga gaccgtgtgg 3600  
 agtgccggat cgacggcatc ggtcagctga ccacgacggc gcgaaaggcc actacacgat 3660  
 gagcactcac gcggtcgagc agctcgaccg gaccgcactc ctgcaatgtg ccgacgacct 3720  
 ccgcgagacc gtcctcgacg acgactcaa gggcgtgcc ctgggtgctc gtgtgctcct 3780  
 cggagagggt gaccagcaga actggaacgt ggcccgcgga gacctcgccc tgccggtgac 3840  
 cacactccac gcacgcgcta tgaacgcaa cctggcgact atggaggcgt actgcgagcg 3900  
 caacaacgcg cttctcgccc cgcacggcaa gaccaccatg tcaccccagc tcttcgcct 3960  
 tcaggcgtct cacggtgcgt gggcgttcac ggcagcgacg cccactcagg tggcgtgat 4020  
 gcgcgcgtac ggggttcagc gcatcatcct ggccaacgaa ctgctcgatg ctcccgcgt 4080  
 gcgttggtg actgccgaaa tcgccagcga cagctcgttc gacttctact gcctggtcga 4140  
 cgatgcggac acggtgcggg cgatgtccga taccatcgcc gactccggcc acggggtcac 4200  
 cctcaacgtc ctctcgagg tcggagtgc tggtggcagg tgtggcgtgc gagacctga 4260  
 ctcggcgctc gaggtcgccg cggcggtcgc cgcgaccgag catctccggc tggtcggcgt 4320  
 cgaggcctac gagggcctcg tgaccggcg catgacaccg gaggacctcc atgcactcga 4380  
 caccttcttc gctggcgctc ggtcgatcgt gctcgagctg gcacgtcgcg acctgttcga 4440

cgccaaccgc atcatcgtga cggcagggcg aagctcgtac ttcgatcgtg tcgtcgccgg 4500  
 actcggctcg tgggacggtg tccgcgacga cgtggaccta gtgcttcgaa gcggctgtta 4560  
 catctcccac gacgccggca agtacgaaaa gctctcgccg ctcgctggcc gacgggccga 4620  
 gcaggaaccg ctccggctcc acaacgcact cactgcatgg gcttcctgtc tgtcccggcc 4680  
 ggagcctgat ctgcgatcc tgcgatcgg gaagcgcgac gcggcccatg acctgacgt 4740  
 gccacagcca cgcgagctgt ttcgcgcgga tggtcgcagg cagcggctgc gggaagcggga 4800  
 gacgttcaag ctgatggacc agcacgcctt cgttcgcgtt gccccgggc tcgcgatcgc 4860  
 tcccggagac atcgtcgctt tcgacatgtc ccaccctgtc acggccttcg acaagctccc 4920  
 cttcatcccg atcatcgacg acgacttcaa cgtcgtcgac ggcgtcctca cgttcttctg 4980  
 acggggctca ggacctcta cgggaggctc cgggtgaacc gacccacaa gctcccgaag 5040  
 tcagttggac atgaccacca gccacagaga ggccccgaca ccgtgtcccg tcaggcgttc 5100  
 acagacgtcg accttctccc cgcgggaggc ccttactccg ccgcgtagc catcggcaac 5160  
 atcgtggcaa tcgcgggca gtgcggctac ctgcgcgaca gatccctcgc accgggtggt 5220  
 ctggaacctc aggtgcgccg ggcattcgag aacctcatga ccgcgctgcg ttcagccggc 5280  
 tgcagcgagg cagacgtgat caccgtgaat gtgttctcgc ctgacgggga cgacttcgac 5340  
 gcgatgaacg cgatctaccg cgagtacttc agcgagcccg gcccgggccg gacgaccatc 5400  
 accgccggtc tccgaccggg tgtcctgttc gaggtcagcg ctgaggccgt ggtaggcggc -5460  
 tgaccgggccc actgccactc ctcccaccaa tgtccctcca tcaacggcgt gccacgactt 5520  
 tccgggcccgc cgtcaccgga aaccgcaagg acatgtcatg gatctcgaa tcaatggaag 5580  
 agtcgccctc gtactcggag ccggcggcgg actcggttcg gccattgccg cctcgctaag 5640  
 ccgggaaggc gcgcgtgtgg cgtcggccga catcgacaag gacgcgctca cggaggtcga 5700  
 ggcgtccatg accggcgata acatgagcct ggtctgggac ttggcgaacc tgcggtcat 5760  
 cgatgagaac gtcaccgccg tcgagagcga gctcggcccc gtggagatcc tggtaacat 5820  
 caccggcggc ccgcgcgcca cccgtgtgtc cgggcaggac gccacctgt ggcgtcggca 5880  
 cttcgaggcg atgatcctct ccgtcgtcgc gatcacagac cgcgtgctcc cgggcgatgac 5940  
 cgagcgggga tggggccgga tcatcacgtc gacttcgtcc ggcgtcatcg ccccatccc 6000  
 caacctcggg ttgtcgaacg cctcgcgtc gagtctggtg ggctggtcca agacactgtc 6060  
 cgggtgagtg gcaagcgacg gcatcaccag caacatcgtc gttccggggc gcatcgccac 6120  
 tgcgcgcacg gccagttgg accaggcaaa ggccgcacgg gagaaccgca ctgcggagaa 6180  
 catcgcgagg gaatcagaga gctctattcc ggtcggccgc tacggggcggc gcgaggaatt 6240  
 cgcgcacgcg gtcacatttc tcgccagcga gcgtgccagc tacatcaccg gatccgttgt 6300  
 ccgtgtcgat ggcggaactg tccagaacgt gtaaccccga caaacatcct gtctccagat 6360  
 cagcgcaacc cgaagccatg tccgtcctc ggggtcactg acgtcagcga tcccagccc 6420  
 aaactcatca cctacgtgcg ctgatctacg gtcattccagg gctaccact cgcagcggg 6480  
 atcacgatca gcagcacgga cctcgcctcc aggggtgcag ctgcgcgcga gcggtgccgt 6540  
 cggttttcgg gtcgcggcg accagcttt tccccacatt gaaagctatc cggaaaccgga 6600  
 tgcgtcctgg aaggcgtact tcagggtatg atgcccgaac gcgaccttca gctacctacg 6660  
 atcgtatctc tggaaagagg tcgtccggtg gcaggagcgc aaacaccggc gtactccctg 6720  
 gaaacagcta cgcgcagct acggcatcca gccagccgac ggcgatgtcg ggctgttcga 6780  
 tccggaaggc cggctcggga gacgtgcagc cgggtgggga aggagcgtg acaaaatcct 6840  
 gacgacctgg cactaccagg acagctcgtg ggcaagcgag atcaggcggg gtatgcgaaa 6900  
 cctccgatct tgctaacagt caagcggcga cccgtttcat ggtcgtgtaa tctcgcggc 6960  
 caccgcgct ggcctcacgt ctgtgcgcca gcggaaatgt cgcacaggcg cggtatccc 7020  
 gaaaggaacc catogaatca tgacgaaccg tcccttcgc tggtcgtcgc tgggcgccag 7080  
 cgacatcgcc cgcgcatttg taggcccagc gatctccgcc cagcctgacg ccgcgggtgt 7140  
 cagcgtcgtc agcggcgacc tggcccgagc gcaggatgta gcggcaccat tcggcgcccc 7200  
 gggccatgac gaactcaccg tggcgcttag cgatcctgac gtcgacgccg tgtacatctc 7260  
 gagtatcaac tcggcacatc acgatcaagc acttgccggc atcgcagccg gcaagcacgt 7320

cctgtgcgaa aaaccgcttg ccctgaccgt cgccgaggca caagaacttg tcgatgccgc 7380  
ggagaaggca ggcgtagtgt tcgcgacaaa ccaccacatg cgcaactccg tgccacatcg 7440  
tttgatccgg gatgcaatcg ccgcccggcg catcggcgag ccggtcgcag tgaacgtccg 7500  
gaacgcgatc cgactgccgc ccgcccgcgc gagatggcgg acgaccgatt cggcagcggg 7560  
cgccggagtc gcactcgacc tcacgggtgca cgacgcggac tgtctccggt tcgtactcaa 7620  
tgccgatccg ctaacggtcg tggcccatac ctctccggc cacatgacga ggggtgggaat 7680  
cgatgagaca atctccggaa gcgcggaggt cgtcagcgggt gtgcatgctt cattcgtgga 7740  
aagcttcgtg accggacatg cccccacatg cctggaggtt ttcggaaccg aaggagcgt 7800  
catcggcacg ggtatccagt cgatgtcacc ggtcggcacg ctgacacacg tcggtccggc 7860  
cgggtccagg aagatcgatc tgggtacccg cgaggacctc tacgtcgtgg gtgttcgacg 7920  
gttcacgcag gctgtgcgtg gcgaggccga gcccccgcc tccggccaag acggggtctg 7980  
gtcgttgcc ttcgcacggg cggcgtgga ggctgctcg accgggcac ggacagactgt 8040  
gggacaggct cgatgacga agctgggcgc agttaggcac aacagatcat gcacaaggac 8100  
aaccgggcat cggcagacgt ggcatcgctc ggggtccggt gaggtgact gatcgtcgc 8160  
ggggcgctcg ctggtcgcgg actcgatgtg ctctgtgttg aacgggacga gtttttggt 8220  
caagagccc ctatcgacgt gtgcaaggct tcgcgcacgg gcactacaag ccggaacgaac 8280  
gggtggtcga tgggtgctgg agaccgttca caccggctac tcactactcg tcggcggcaa 8340  
tcagccaaggc ctacgaagct ggcccgtggc tcaatctata ccgcgtaagc tcgccaacgg 8400  
gtcctatgtg ggccggccga actacatgca gcacccgaac tccgcggtca tggccatcgg 8460  
caatggccag cctccgatca tggaccatac ctacttcaac atgcgccgtt gcgggggccc 8520  
cggacgatta gctgaagcgg atgtccgctt tcggtatcgc agaccttgc acgacgggcg 8580  
gtgttgccgg gccgcgaac gatcatgtga gccgggctgg cccgagaagg gcattggcca 8640  
acagagcagg tgactacgtt atctctgccc ggtcatccaa ctgatgacct gtgcgcgcga 8700  
ggtaaagccc agcttggaag ggatgttctc gacgtgggccc tcggctgtgc gctgggaaat 8760  
gaccaggcca gccgcaatct ctttgtttgt tttgccagcg gtgaccaggg cagcgacctc 8820  
gcgttctcgt tttgtgagaa ccggttcggt ttccggagcc gaagcggcgg ccgcttcgac 8880  
cggcgggtggc ttctcccaaa gagcatagaa gatagtctgc tcccggttca tccgcgtccc 8940  
ctggcctcgg gcggtttcgt attcggtgtc tgtgagtgtc tcctgggcgc gtgccttggc 9000  
ctgctcacgc caggtcagta gcgtctcgga tccgaacagc ggggtgccaa tcagttccca 9060  
cattttatcc aacgccccga aaaggatcgc cgcgcgttca ggattgcgtt ctgcggcggc 9120  
gacccaggcc agcagttcaa cgcagtcgct gatcccgagt aggtcgttga gttcccgtt 9180  
attgcggagg gacctggacg cggaggcggc tgccttggtt gggtcgccc cggcccacca 9240  
agtgatcgcc atgttccact ccgcccacga tagcctccag cgctcgccaa gttctgtaca 9300  
gatcttgccg cattcgtgca gaaggacat cgcccggtca aggtctccaa gcaggccgc 9360  
tgcttgcgtc cgtattgaga aaatggctaa ggccgggccc gtccagtggc cagttgcgcg 9420  
gtggcgggccc agtgccctcgt cgagcagttg cgtggcgcga gtcaggttgt tgtcccat 9480  
ctcccggtcg ccaagaaact gagtcgcgta ggtaaattcc gtttcgtcgc ccagtttcg 9540  
ggccagctcg cggctttcgt taagcaggac cagcgcceca tctcggtccc cttgccggtg 9600  
ggcaatccat ccggtgatcc acagtgcgcg ggcacgctgg ctggtgggtt cggctccag 9660  
ggccaacgcc ctgtccagcc acatgcggcc ctccctcaga tagtcgccag cagcccagta 9720  
gaaccacagc gcgctcacca tccgcaaccc aaccgacgcc tcgccagggtg tggatgaagca 9780  
gtagtccaac gctgccaga agttgccgcg ctccggcgcg agacgcgcga gccaacggc 9840  
ttgatgcggg cccgcggact cggcatcgga ctgctcggcg agatgcaggt agtagtcgcg 9900  
gtgccggcag cgcagagcgt tttcggcacc cgcctcgggt agccggctcgt gcccgattg 9960  
acggatagtc tccagcatcc ggtatctcgt ctctcgtcgc tctcgggtcc gggtagcac 10020  
ggacttgctg atcagtcctc cgagagcggc gaagacgtca tcacgggtca gcccgctgcc 10080  
gggtgcagacg ttttcggctg cgtcgagatc gaattcgtcg gtgaaaaccg aacagcgcgc 10140  
ccacaactgc ttctcatgct cattgcacag gttgtagctc cagtcgacgg cggcgcgcag 10200

ggtctgatgc cgaggatcgt cggcgcggtc gccggtggtg agcagccgga atcggttctc 10260  
 cattcgcgtt agaatttgcc cgactgagag cagcgaacg cgcacggcgg cgagttcgat 10320  
 cgctagggga atgcggtcca gcctccggca taacctggcc acgatctgtt cgttttcttc 10380  
 gttgatggtg aacctggta gcgcggcgc ggacgctgc tcgaacaagg tcaacgcttc 10440  
 acgctggtag ggacgacctt cctgttgta cgttgatgaa tcgctgccag caggcacgga 10500  
 cagtggcggg accggccagg tctgttcgc cactatccct aaatgttctc ggctggttgc 10560  
 aagtatccgc aaccgcgggg ccgcgccag cagcacagcg actaggcggc tgcaggcccc 10620  
 cagcaggtgc tcgcagttgt ccaacaccag caagacttgc ttgtccgcca agtaatcgac 10680  
 gagcacggcc tccggggcac gtgtcgaacg gtcattgatc ccaagagcag ccgcgaccgc 10740  
 ctgcggcacc agcatgggat cgcgaacgtt agccagctcc accagaaatg tcccgcgga 10800  
 gaaggccgc cgcacgtcgt gtgcgacgtg cagcgccagc cgggacttcc cgatcccgcc 10860  
 gaaaccagtc agggctacta accgtgaact ggacagcgcg cgcttaactc ttgtggcgc 10920  
 ctgacggcgc cccacaaagc tcgtcacctc atatgggagg cgaccagcct cccgaaccgc 10980  
 aacccagac acgactcacc gcacctcggg agaactagcg ttttattgta gcccggtcat 11040  
 ggctggcagc gcacttacag cttgaacaac tgttcgaacc acttcgaacc caccgacggg 11100  
 gttcgttacg atctcggcgg atgttgatcg tgggtgggtgc ggggagggcg tgcgttgc 11160  
 gctggcctgc ccagcctgta ccactgatca aggtcgccga ccgggcccgt gtgggcgatg 11220  
 ttggcactgg tcaggccact gctggcggtt atgcggcggg ccagctgcgg cggcacggtg 11280  
 ggcggcatcc gctccgctac tgggtgcgca aggcgctgc gtccgcctga cggcgaaagg 11340  
 tgagctcgtc gtggaggcag cggaccgggt gatcgctgag ctggagacgc tgcgactca 11400  
 gatcaatcgc tcacgtggcg agatctcgca acgtctgcgt gtggcctcca tgcagaccgt 11460  
 caccttcac cgtttccccg aggtgctgac tcgtctacgc cggcggcgtc caggactgga 11520  
 tgtgtgatt gttcaagccg agccggatgc ggctcttcag ggtctccagg tcgacgacta 11580  
 cgacctagt atcgccgagt cctacccgca gtggccggtc atctatccgg accggctcac 11640  
 gacagagatc ctgatgaccg accgcatgca cgtcgccgca gctcccattg tgggtccacc 11700  
 gcccgctcgc tgggggcagg cgttttcgcc gctggcgagt cgttcaccgc tctcggcgct 11760  
 ccaacgtctt gcctgctcga agttgtcccg agtcgggccc gacgatggct gtggtgaagg 11820  
 gatcttgacg ctcatgcatg gaggttcacc attgttgagg caagaggctg atgacggtg 11880  
 cggcggtcgt tgcgccagcg gcgatcatcc agccgtgtgg gtagtagcgg ctggctccga 11940  
 agaattggct caggccggga gtttgacga cagcgccag tggcacgaac gagacaagcg 12000  
 cacacgcat gaccagtggg gtccgtccgc gcacgaccag ggtctgtccg agctgggcgg 12060  
 cgaccagcgc cagcagcgcg accgtgttg cctggcgag ggttcgggtc gccctgcccc 12120  
 ggatccacgc acaaatggcg gcccgggcg tggccgggc cggcggttag atgtcgcg 12180  
 tcagcgcagc gccagcgag gcttcggggc cttctgccag cagcatctcc gcagtgatcg 12240  
 tcgggggtg tgcaccgcg atcgccatcg ccgggaggat gtcggtgagc aggttgacca 12300  
 gcagcaactg gcgggcgttg aggtgcccc cgcggtgat cacaccggtg acgacggtga 12360  
 agccgatctc gccgaggttg ccgcgagca ggatggccag cgagtcacgc accgatgcc 12420  
 acatggcgcg gccttcacac atggcgctga cgatggtctc gatgcggtca tcggtgacca 12480  
 cgaggtcggc ggcctcgcg ggcgggggg ttgcgcgctc gccagcgcg atgccgatgt 12540  
 cggccaaccg gatcgccggg ggcgtgttg cgcgctccc ggtcaccgcg accacacggt 12600  
 cgttgcttg caggcgctcg acgatgcgcg cctctgcgc cgggctgacc ctgcgaaca 12660  
 ccgcgacggt ggtcacttcc gccgcagtt ggtcgccgc gaggggtgac agttcggtgc 12720  
 cggtcatgac gcgacgtcg ttgagcgct cgagttcggc ggcgatcgcc tcggcggtgc 12780  
 tgggtggtc gccggtgacc atcacgacct cgacacctgc ctgctggagc tgcgcgacg 12840  
 cggcgccgc agtgggcgc accgggtcgg ccagcgcgac gaaaccgatc agccgagat 12900  
 cccgatccg ggaactcgtc agatcactgc ggtccgaggc ggcccgctcc gcgacagcca 12960  
 gcaaccggta gccctgcccg gcgaggttt cgacaactcg ctgcagcct cgacgagcgg 13020  
 tctctccaa cggcacgtcg ccatctgatc tgcgccacc ggtgactgg gccagcacga 13080



tctccggggc acctttcaca ctcagcagat tgccgtcgcg ccaggtgccg agcaccgcgt 13140  
ggtagccacg ggaaggctcg aagggcatct cgtcgatcgg taccagcgc tgettgcctt 13200  
cggccgggtg cactcccgcg ccacgcgcgc catcgacgac cgcgcgatcg gtgggatgcg 13260  
ggagcagttc tccgtttctc accaacggac ttgcccgcaa ggcaccgcg agcactcgcc 13320  
gttcactgt ggagatttcc tcgagccacc ggctcgacgc gccatcggat acctgccga 13380  
gcgcgatgcg ccttcgggtg aggggtgccg tcttgtcgaa acacaacacc tgcgcccgcc 13440  
ccaaagcctc gatcgtcgcg gagttacgca ccagcgcccc gcgtttcgac agccgcccgg 13500  
cggccgcgag ttcggcgacg gtggcaacga acggcaatcc ctcggaacg gccgccaccg 13560  
ccaggctcac cgcccccccc agcgatgtcc ccaatggacg cctacgcagc atgtccgtga 13620  
cgagcagaat cccaccgcgt cccacgcaga tcggcaccgt caccttggcc agcgtctgca 13680  
ggcgagccgc gacaccatcg gcacctcgcc gggcatcggg cccgatcgcg ctgcctgcct 13740  
cgggtgttcg gccggtcgcg accacgacac ccaccgcccg cccggcagcg accgcggtcc 13800  
cccggtagat catgcacgtc cggtcgcga cggcaccgc cggagtcgcc gtcaccgtct 13860  
tggtgaccaa catcgactcc ccggtcaggc tcgactcatc gacctccagc ccgtcggcct 13920  
ccagaagccg gcagtcggcc ggaacagcat ccccgccac caactccacc acatcgccga 13980  
gcacgagctc gtcagccggc aactgcgtca cattttcctt acggcgacc acaaccgaa 14040  
caacgctggc ctccagaagc ctgttcaaag cgcggtccgc ggtgatgcg tgcacgccgc 14100  
cgatcaacgc gttcatcccc agcacgccc tgatgatcgt cgcgtccgca acggcaccga 14160  
tactcgccga gatcgcgcc ccaccgcca acgcccgcgt caacgggctc gccagtcct 14220  
ccagcgacgc ccgcaacaca cccatcggc gaaggccgt cccctcgtec tgaccttga 14280  
aacgctcgtg cgcgcgatca tcggtcagac cctgcgatga actgccgagc agctccagca 14340  
ccatgccggg gcccatcgca tgccaggcag tgcgcggcac cgcgcgcgga accggacgcc 14400  
acgcccgtc catccccac cagtcccag cgcgcgagcg gattgccgca cacacggcca 14460  
ccggaacgcc cgcacggacc tggctcccga tggcaggccc gaacgcgcg aacaccgcc 14520  
cgctgaccga accgatcaac gacagctgcg cgtacgcac gctgacataa cgcgcggtcg 14580  
ggaccgtact caacaacgca cagactccc ccaggctcgc gcccagatc aggtgcgcac 14640  
cccacggaac ctaccacca cgcgtcacac caagcccgc atccgcagcc gccagtgcg 14700  
ccgcgccccg cgcggacacc acaccgacca cgtgccctc ccgtgcate tcccgacca 14760  
ccgaagccag ccgaccccca cccgggacca cgcgatcaac accgatccgc cccgccaacc 14820  
ggccccatc acccgcgaga accaccgacc cgacctcgcg agtgccgcg accaactcct 14880  
cagcgagcgg atccagctcc tcggcagcca gaaccacgc ggccgtgcgc gaatcgtggc 14940  
gcagcgcgaa cacctcgata ccccgaccgg ccacctcgcg cagctctggg ggcaaccgcg 15000  
accgagcgac cgagtacagc gaccaaccat cagctcgcgc gccgacccc ttggtcacac 15060  
ccacgagttc atgcgcacgt tcccacaact cgacgaccgc gcgcccccg tccgcgga 15120  
tcacgtcatg caccaaccga gcaccggtgc gcagaacgtc cgcgtcgatc accaccgtgt 15180  
cgaccgggtc caaccgccc aacaccggg catcgaacac cagagtcgtc cgtccgaca 15240  
gcccgtaccc gagctgcgcc gcgaacgcgc cccgccccac ctgcgcgcg cgcggcgcgc 15300  
ccgcgaccac catcgctgc gcacgtgaa aactccgct caccgcgagc atgcccgcga 15360  
acgcccggg cccaccaga gccgaaccgt tcgcgcgcg ctcgaccggg ccggcgggca 15420  
acggcacagg acgcggccga acccccaccg gcgcactaog gtgactggca ggatgcgcc 15480  
ccgttccggc tcccacgc gccagctcc gcgcgcgc ctgcgtctcc cgcaacaggc 15540  
agaaccggta cccgattcc gccagcaaac tcgcccgaag ccgcgcaagc gtgttcccca 15600  
acgactgct aaaccgaac aacaggctag tggcggtctt cccgatccgc gtttccagcg 15660  
gccccggac ccaggaggcc gactaccgg tcgacaccag cgcgggaaca gacggcgga 15720  
gcacgggcag cggcagatc cgggccccca ccgcgatcac agcgcgggtc aagttcgac 15780  
ccagcaccgt cgcctcagcg atcaccttga ccgattgtc cgggtgggtc acaccgaccg 15840  
cagcgaacgg ttcctcgtec aaaccggtgt ccgcctccgc gtcctccacc acaccacca 15900  
gctcagcgac cccaaccagg cccggatcgt gcccgatgac caccgcccc acgacagcgt 15960

tgatctcgac ccacgcaaca cccggcacct tctccaaccg cggggcgatc tgctcgccag 16020  
 cctcctcagc ccccggtctg tcaacaccgc gcacctcgac gtacgcacgc cctcccgacg 16080  
 accacgcccg cgcggcgcg cccaagcccg cgatgctctg cgcgacagcc accaccggcg 16140  
 ccacgatctg cgaggccaac gcgcccgaag tacgcagcat cgaacacact cagctgtccg 16200  
 acgaaaagcc gaccgacgcc cgtgacaccc ccgcgcgggt cgcccccgca accggccatt 16260  
 cgagccccc aaccgaccgt accgcgcgaa ccgcgtcctc ccaacgaccc ggcattgcgaa 16320  
 catcgaccgg acgaaactgt gccacgacgg gcaggctgac cgccgcgggt cgcaaccgcg 16380  
 ccccttggtg ccgcgcgcct gcgccaggcg ttgtgttccg cgccgtccta gccaccacta 16440  
 cctcccgcg gcacagatac cgtgagtcac gagcgtagga ccgacgagcg gaaagcgaac 16500  
 ctcgaaacag ccgcacatcc gtctctctcg tggcctgcga gctgatcagc aacaccgacg 16560  
 gcttcgccac gotgttcgac ccgccgacac accagtagtc ccagcgggtg tgaattcaag 16620  
 gccgcctgtg tgctcggggt cttcgaaacg gctccgagca tgcagctacg ccggacgcag 16680  
 gcaacgcaac acctaccag gttcacctgg cctgagcgcc gccgcggcg gctcagcgac 16740  
 gcctagcccg gaaccgcgta caaacccgat cacttcgacg ccgggtccgc ggcacgctgc 16800  
 gcgaccgga caccacgccc aatcgccgcc gccaccggcc agtccatcaa ccccaacagg 16860  
 gtcagcgctc caagaccag atagaaggcc atcgtctgat ggtccgggag aacgccccga 16920  
 acggcctcag ccgcgctagc aagatcatca cgagtgggaa tgtagtgatc gggacgggtg-16980  
 aactcgccg ttacgcccgg cagactcacc gtggttcccc cagcaccgcg cggtttccgc 17040  
 tgcgtagtaa ccacggcacc tctcaaacac agggcccatg ccgcataatc ggccctgcac 17100  
 ggactcaaaa ccgcccgaac cgaccaaccg ggcaacacga ttttatctct gaacacgaac 17160  
 gtccccgcgc cgagacgggt ggaatacggg ctgcgcgcgg acctgaacca gtcaggcggt 17220  
 gtgcttcaac aacgttgctg tgcaccacc tgctcggggt ggccaaccgt gtggacgcaa 17280  
 cgccggctcc agtgactgct ggtgatcaca ctcggtgate atgatcgtct ggactgcggt 17340  
 ttgatcaagg caatgttccc gcacctggcc ggggtgcggg agcggctggt gcgcgggttc 17400  
 gcggagtcca ccaacgagta tccgtggcgg tggacaccgt cgcattgtga cgaatggtcg 17460  
 caatcacoga ccggagaacg gcatctcgcg ccatccaca tccgcgccta ccaaggcagt 17520  
 ctgaggctgt tcaacgaatt cctctgtgac agccgggtat gctgggcgct ggcgtgcgag 17580  
 gcgtcgttcg ggccggacga gtaccgggtg gcgaatgtgc acgagtggaa ctactcgcc 17640  
 catctccagt cctatgaggg cgatccggac gaccgcccgt tcaccgcga tgagctgcag 17700  
 cggttccttg actacgccac gaccaagtgc accgcgcggt tcgcgcgaaa cgcaacggcg 17760  
 cattggcagc ctatcgagac gcgacctgt tcaaagtgat gtatcgctag ggcttggttcg 17820  
 ccggggccct acccgcgctg gctcccgcga ttttttgagt ttgtgcagtt cagaggacga 17880  
 tgtgggagat gctagccgtc gcgacagatg tgggtcaaagt aggcgcggca ggcaaacctc 17940  
 agcggctcgc gtgcagtatg cggctcgcgc gcagtgcgcg ggggtggttcg cctcgaaac 18000  
 ctgcggatcc gtggaggtag ccggcgatga ggtccggggc ggagtggctc tcgatgccgg 18060  
 tgaaccgag ggcgtgaagc tgtcggcgga tgtcgtcgggt tgtgaagtag ctgaacaagg 18120  
 gttcgccggc gccggccagc cgatcggcac gtgcccgtg gtgcgcacgg tctcgtcgg 18180  
 tggccgcagg ctgcaggtag tcgaagacca cctcgaccgg ctgggcttga ccagcgatgt 18240  
 attcgagggt ggcgtggcg gcgttcgcgg tcagatagaa gacaacgccg agccacacga 18300  
 atactgccgg gtccgtccgt ttgaatccgg cggattccaa tccagttgcc agcggctggg 18360  
 tttcgaaatc gaccggtacg aaggtcaacg tctcagggcg gtcgatccct gcgctagcaa 18420  
 ggcgttcgcg tttccatgct tgggtggcg ggtggtcgac ctcgaaagac cgcaagtccg 18480  
 ggcgcggatt gcgataggcg aaggtgtcca ggccccgcgc gaggatcacg acctgccgta 18540  
 caccatcggc gatggccgc gccacggcgt cctcggcgaa gcgggcacgg gcggcgaaaga 18600  
 acaggcggcg gggctgatcg attgccccg cgccgaggtg atcggttggt ggcgtggcca 18660  
 atttggtcag tctctctgc gtgacgccga gcagacgtac cgccaggggg tcggtgaaga 18720  
 tctgtggtcg gtcagcgatc tgggtgatag cgcgggcgta ccgggtagcg agcgtgtcc 18780  
 ggtgggtcc gccgttctcc atgttcaaga acttactcac gaacgcggcg cgggtctat 18840

gaactcactc acgcccgctt cacaaaggag cagcgaaca cgatccgtaa ccacagctgg 18900  
 ccctccttta gggcgattgc cagcagtcga caggcagggc tcgaaccctg ttatcggggc 18960  
 tgggttgatg tggtgagccg ggtcatgagt tcgcggttgg tggccccggc gcggaagat 19020  
 cgccggcgcg ttcgtcgagt tgctgttgct gcggcaagaa acgacagctc ggccgaccgt 19080  
 tcaacggtea cgggtgcgttc ggaccgcgtc ctgccaccac acctgacggt ctttgcgtc 19140  
 accgggccct cccattgat gtggccgctg atgctcgaat gcgcacgagt ggtcagtcaa 19200  
 cgccggttcc cgcagcccg cttccccccg gaagacggcg cgttcccgca cggcagcccc 19260  
 acgcgttcac ggggacgata atcgggtccac gaggcgaata cctgctccgt gcaggtggaa 19320  
 gcggctattc cagggggccc gccgtgggtg ctgcaatcgg cgtcgcacg cggctgacgg 19380  
 tgggcgcggc atgatgagcg accgtgctgc tgcgactggc cttgectatg ccgataacct 19440  
 cgtccgtgca ggtcacacag cagcgtctga cggttttccg gtcgagcccg gccagccggg 19500  
 tcacctcagc cagcgaatgg cgcgccagcc acagctggct tggttctcag cggttgaaca 19560  
 ccggtagctc ggctggagggt tgcgggatga atccgccgtg ctggggccatt tggctgtaga 19620  
 gatgaccgtg cgactccgag agcgttgcaa gcgctccgtt gacgtcgacg tgcaccagcc 19680  
 ggccgttccg cttgaccacc cgcccatcga ccagcaccgt gtgcacgttg cccgagttgg 19740  
 tctgcgcaat gatcgcgcca caagggtcgt ggcgggtcca accggcctgg ctgatgcctg 19800  
 acatgtccag caatacgatg tcggcccgtc tgcccgggtg cagggttccct gtcacgtcgt -19860  
 cgacgcccgc ggcgcgggca ccgttgaccg tgagccaccg taaggcgtcg cgcgttttcc 19920  
 accgcatgac cgtgggcagt gtcgaccgcg cgtagtcttc ttgatcgtcc ctccatcggg 19980  
 tcacctgtag gacgagccgt gcgtgcgaaa tcatgtcgcc gcaggtactg ctgctgcagt 20040  
 cgatgccgag gcttgcccc ggcggttctt cgggtggcctc acgaatgacc gggaatccca 20100  
 tgcccatctg catctcggtc tcggcgcaaa ccgagatcat ggtgccggta ccccgagca 20160  
 ggtgccattc gttctcagt ctgaatgtcc cgtgcaccaa gagaaggctc ctgccagca 20220  
 tggtgtggga gtgcagcgc tcgatgtcct tgaacagggt gcgcgcgatg acttggttgc 20280  
 agtgaacgt gatgcgtgc ccgagctcgc gggccagctc gaattcctgc ttgacgtccc 20340  
 cgaagggagc gatggctagt tcctgcgggg cgatgccgaa gcgcagcaat tgttgctcgc 20400  
 tcgggaagta ctgggcgcgg atttcccggg ccagccgcgc ccgcccggcg aagtcgcctt 20460  
 cttccaccga ctgctccagc ccgcctttgg tctcggaacca cgtgttcgac gtgatgggca 20520  
 ggaggccgtg gccgtacagc gcgcgcacgc ccgcgtcgcg gaggccagcg accgcccgt 20580  
 gcgcgtgctc gccggtgacg atgttggtgc agtaatcgac cagcgtggtc acccccgagt 20640  
 taaggcaatc gaggccaccg aggtagttgc ccgcatacat gtccctgcggg cggtaactcg 20700  
 tcgccatctg cagccggaat cccctcaggt agtcgaggat gttcccgtcc gccagaatgc 20760  
 cccggagccc ggtttgccag gtgtgtcgat ggtgtcgac catgcccggc atggcgatca 20820  
 tcgacgaggg gtcgattcgc tcggtgtcgc caacgttgag gttcggcccc atctcgacga 20880  
 tgacaccgtt gtcgacgagg atgtcggctc gttcgaagtc gccgagaacg tegtccatcg 20940  
 tcaccacggc agcgttctcg atcagtgtgc gcattggcaa gtgctcctct ccgcgcggtg 21000  
 ggcacgcat ggcgtcctca agcgggtccc ttcggtcgtt gatcgactgt ggagaagagg 21060  
 ctgtggagag gggccgtctc gcagacgggg cagcttctac tgggtccacgc caatccgaaa 21120  
 accggcggtt gtcagtcgca tcgtcgcgcg gtcttctgtt gaatgacca ggtcgcgatt 21180  
 tgctcgcgag tgggtgaagc gagtttggtc aggtgtgtct cgacgtgggc ctccagcgtg 21240  
 cggcgggaga tcaccatgat ggccggcatt tctttgttgg tgggtgcccgt tgccagcagc 21300  
 gctgctacct ctgcctcgcg cggagtaagc accgacgcgt gcgtttcttc cgtcctggag 21360  
 cttgctgggt ccccggtttt gcccgattc cgtatccatg cgatggcctc gtcgaagtcg 21420  
 agctgtgtgg cgtggtgaac gcctttctga aacgtctcgt cgcagagtga gctacgtgct 21480  
 actgcctcgt agtggctcgt taggcgcgca acatgtttgt ggccggcgag cgacgcacgc 21540  
 acgttctttt tgatcgtttg ggacacgcg aacaaggctc cagcccgtgc cgggtcaccg 21600  
 tctgcgctgg ctatgcaggc cagaatctca acgcacaacg cgatctgaaa ctgctcatgg 21660  
 aagccgagcc gcaggcgaag actctccagt tccgcagcat tggcttgccg gtaattgccc 21720

gctagtaggt ctgctattcc ctgcgccac agggcaagcg atctcatcca ctgttcgctg 21780  
taggtctcct tgacggcagg gcacccgcgg aacagggcag cggcgcgctc atgatcacca 21840  
aggaacgcga ctgccaacgc gagctcgatc tgatcgtagg cggcggcagc gtgatcgcca 21900  
atcagccgat gcttggcgag agcgtcttcg aagagcgcgc tcgcccgcgc cggatcttcg 21960  
ttgaacaggg cggtagccc cgagatctgc gcgacgtagg ccgcccctga ttggtcgccc 22020  
aactccgccg ccagcgcgct cgcctgctcg agaagggtg acccggatga cgcgtcggtg 22080  
agcgcgcat tcaggtaccc gtgcagctag agagctttga cacgtaccgg atccgcacgc 22140  
aatcgatggt aaccggacat atgoggctaa tgggtgcacaa cgggcagtgt tgatctgacc 22200  
acgctaaccg gtatctcggg cgggtcgctg gctgcgccgt tcaagctcaa gcgtgcgggtg 22260  
cgtcttccgc tcgacgatct acaagtacgt ccccgaaactc agcaccacc gagcgctcgc 22320  
ctctgggccc ggtccgcagc cagcgttcca gagctaagcg ctggtttccg gcgtctgcta 22380  
ccgggacgct tcggccgcgg tgggttcagc gcaagtgagg gttcctcggc ggaacccaa 22440  
gggtccgtcc cggccaggac caggactgga cccgtgggtt cggctcaacc cggagtgtcg 22500  
catttgatcg tggtttcag cagtaatgga tgctgaatct tgtgattgac gccgccgcga 22560  
ccggagagtg tcttgagca accgattccc gcgcgattcc ggaccatcca tcggacctag 22620  
gttctgttgc gacctgagcc cgcggcgcg gcggagtttc cgcctcaagt ggccagggtcc 22680  
agccattcct cgatgggcag cacgtgtgcg cgggtgtccg cgtgcgtgcg gaactcattc 22740  
tcgggtcgg agacctcgac gtagcccgcg atcttgctcg gctgttcgc ctcgtagtgc 22800  
atcggcactg cgtaccgggc gtgcaggatc tctgcggctg cggcggcctg cctcgggtcc 22860  
atcgcggcg gcagcgggct cggcggtg aggtgtggcg cgtcgaccac cgcgccgttg 22920  
gcgggcagga acaccgcgtc gaacgggctg aaccggcgcg cgatgagcca ccagtagccg 22980  
tggaacatcg tgcgcgcgc gtggaagacc cgtgcccgt cggcctgcac caccagttc 23040  
agctgcgggt cggccagccc gtgcagggc gggaccgcgg tgacgcggaa cggcccagg 23100  
tcgcgggtg accaggcatc cagcacctcg gcagccagtc ggtgcagtgt cagctcgcgc 23160  
tcggccggca gtgtcgtcac gttgtccac tcgtcgccgt ggccgggcgc gggtcggagc 23220  
accggtgccc ccggtgtcag cgcggccgc agcgcggccg cgtcgggtgt gtcccgggtg 23280  
aggtgggtga ccagtgcggc ggtgaccgtc ccgctcggca ccgccagcct ctcgccgggt 23340  
ttccacgccg tgaacagcg tgagaggtcc cgtacgtagt cgatcaccag tcgcttgccg 23400  
ccgcctcga tttccagtcc ggcccagccc agtcgtcgca cccgatctc gcactccttc 23460  
ggtccgtgg tgaagcgga tttagcgtac gcaggttcac taaatcaata gcgaacgcgc 23520  
gtacgtatc ctcgtcacat gtcaccgcga cgctcagcg ccgaagcaca ggccaccagg 23580  
agccgatcc tcgaccgtgc cgcgagatc gcatccgagg aagggtgga cggcatcacc 23640  
atcggccggc tcgcccagga actggagatg agcaagtccg ggtgcacaa gcacttcggc 23700  
accaaggaga cactgcagat ctccacgtg gacaaggcat tcgtggactt ctggcaccg 23760  
gtggtcgagc ccgactggc cgagccgcg ggtctgcggc ggctgcgcgc ggtgtgcgc 23820  
aacgccgtg gctacctaga agcgcactg ctgcccggcg gctgcctgat gaccgcggcg 23880  
ctcaccgagt acgacggcg ccccgccgg gtccgcgac cgggtggccga ggtgtggtcg 23940  
cgctggcggg gacagctgc ggcggaact accgcggcg tggagaacgg cgagctgccc 24000  
gccgggttcg acatcgacca ggcgtgttc gagatcgtc ccaccgggt ggactgaac 24060  
gcggccatgc agctccagca cgaccggcg gccgcggcc gggcccgcc cgcgatcgaa 24120  
cgggccttg accagtctg acgcccgcgc cagacaaacc cggccgatcc tttcccgagg 24180  
accagaccgt ctgaccgct tctgtgtgt cagcccgtat ccattccatt tttcggcgag 24240  
tacgccctag catggtcacg agacattccc ggagccgtca ttgattgact ggcgacgca 24300  
aaagctact gccggttgc ttgcacctg ggtcagcttc gggagtccg ggccagagt 24360  
ccggtgggt cagtgtgagt gaaggttct cgtgggagc ccgaggggga ctgccaccag 24420  
ggtctgaaga cttcgtgaag aaggtatgaa cccgggttg gcatgcttc ttatgaggt 24480  
atgggaagcc aaggcttgg agcgtggaca tcatggcaa cgatcacggc acggtgctcg 24540  
ccaccatgat cttcaacct gtagccccg cgggtgtgat cgcgccagtc ggcgtggagc 24600

tgcggtacga	cagccgcaat	ccgtacgaga	tctccatgaa	gctcaacgta	ggcacggacg	24660
gtcaggtgga	ctgggtgac	gcccgcgacc	tgctggccga	cgggctgac	gccgaggcag	24720
gcgaaggcga	tgtgcggac	ggccctcgac	ggggttttcc	ggggttggtc	gtgatcgaga	24780
tgagctcgcc	gtcggggcag	gcctccttcg	aggtgaatgc	tgaccagctt	gcggacttct	24840
tgaacgacac	ctacgacgtg	gtcgaacctg	gtgatgaaca	ccggtggatg	aacgtcgacg	24900
aggtgctgag	ccagctgctc	tcgcacaacc	tgtaatggcc	cagctctccc	gaagcgccgc	24960
acgccaaagc	gctgggtgcg	ggacctggcg	gcgctgaaca	ccgccacgct	gtgtctccga	25020
gctccagctg	gaccacgtcg	gtgccgtgcg	cccggctcgg	tcaggccgaa	ggtgctgac	25080
ttctccaggc	gcgccatcgg	cgcaggaagc	gctgcttctg	ctcccgcgc	agtaccgtcg	25140
tgtcatggcc	acggacagct	tcgattcctc	gaagctacag	gcggccgtgg	catcgagcgt	25200
cgcgtcgtgc	gtctcggaag	tcagccgaga	cgtctacacg	cacctgatta	ccgaggctcc	25260
gcagttgcga	gccgatgaga	tcgtcctcag	cattctacgg	acgagtgttg	agggaaatat	25320
cgccacattg	ccgcacgttc	tcgaattcga	gattccgttg	ggatattcgc	cgggtcctgc	25380
tgcggtgttg	gagtatccgc	gacgactggc	gaaacatttc	catcaacgcg	ctgatcaggg	25440
ccaaccgcat	cgggcacttc	cgttcctgt	agtgatgcct	cgacgagatc	cgccgccaat	25500
gcgcccagca	ggccgtatcc	gcagcgacca	cgcaacgaat	gctcgcaacc	agcttcggct	25560
acatcgaccg	cgtcacggag	cagatcgccg	aaacctacca	gctcgaacgg	gaccgctggc	25620
tcctggcgac	gggacggccg	tgaggtctct	gcggcatccg	catagcgtct	tctcccgtg	25680
aggcacatga	ggtgttgcg	gcggtcgttt	ccggcagtcg	cacggcattc	gtcctagctg	25740
cgggcaattg	agggagcgaa	gatttagagg	agtgtggcca	cgcggaccaa	gccggcgagt	25800
gctcgggagc	ggctgtgggg	cggccaggcg	atgactgtcg	tcacgtccgg	cgctctaga	25860
accggtacgg	cggcgaggcc	ttcgagcagg	ttgacgcgac	tggattcggg	catgaccacg	25920
gtagtgccgc	cgagtgcgat	catttggaac	agttgcgtct	ggttgcgtag	ttccacgccg	25980
gggccatctg	gatagacgcc	gtcggggccg	ggccagcgcg	caagcgggag	atccggcagt	26040
gagctgacat	ccgccatccg	tacatggggc	tcgctggcaa	gcggatgcga	ggtcgggaag	26100
atggcgactt	gttgctcgtt	gttcagaatt	tcgatgtcga	gttcggccgt	cgggtcgaag	26160
ggttgatgca	acagcgccac	gtcggcccgg	ccgtcatgca	gcgttttctg	gggctgggat	26220
tcgcagagca	gcaggtcgac	ggccacggct	cccggctcgg	cggcgtacgc	gtcgagcaac	26280
ttcgccagca	gctcaccgga	ggcgccggcc	ttggcagcca	ggactagcga	gggctggctc	26340
gtcgcggcac	gctgggtgcg	tcgctcggct	gctgccagcg	cgccgaggat	cgcccgccct	26400
tcggtcagca	gcattgcccc	ggcttcgggt	agcgagactt	tcgggctggg	gcgttgacgc	26460
aacacgaactc	cgagtcgttg	ctcgagctgg	gcgatcgtcc	gcgacagcgg	cggctggggc	26520
atgcccaggc	gctggggcgg	ccggccgaag	tgcaactcct	cggcgactgc	aacgaagtac	26580
cgcaactccc	gcgtctccat	ccgtcgagcc	taccgctgat	tcatatcagc	tgggtatcgg	26640
tgtgagacct	agatggtgtt	ggttccccgc	cggtttcggg	ccacgctaga	aagcatgagc	26700
gaacagacga	ttgcactggt	caccggcgca	aacaagggaa	tcggatacga	gatcgcggcc	26760
gggctcggcg	cgtcggggtg	gagcgtcgga	atcggggcac	gggaccacca	gcgcggggag	26820
gatgccgtgg	cgaaattgcg	tcgggacggc	gtcgatgcgt	tcgcgggtatc	cctggacgtg	26880
acagacgacg	cgagcgtcgc	ggctgctgcg	gctctgctcg	aggagcgcgc	cggccggctc	26940
gatgtgctgg	ttaataacgc	cggcatcgcc	ggggcatggc	cggaggagcc	ctcgaccgtc	27000
acaccggcga	gcctccgggc	ggtggtggag	accaacgtga	tcggcgctcg	tcgggttacc	27060
aacgctatgc	tgccgttgct	acgccgctcc	gagcgcccgc	ggatcgtaaa	ccagtcagc	27120
cacgtcgctt	ccctgacctt	gcaaaccacg	ccgggcgtcg	acctcggcgg	gatcagcgga	27180
gcctactcac	cgtcgaagac	gttctcaaac	gcgatcacca	tccagtacgc	caaggaaactc	27240
agcgatacca	acatcaaaat	caacaacgcc	tgccccggct	acgtcgcgac	cgaccttaac	27300
ggcttccacg	gaaccagcac	gccggcagac	ggtgccagga	tcgccattcg	gctcgccacg	27360
ctgccagacg	acggcccgcg	cggaggcatg	ttcgacgacg	ccgggaatgt	gcctggtga	27420
ggcgctcagt	cggcgatggt	gcaatcgaag	tcggagaggc	tcgctgcgac	cgggtacgcc	27480

gaacaacacc tgttccctgtg ggtacggatg tcggccttcg ccgtctcggg cattgacaac 27540  
 ctgtacttcg ggcgcggtta ccgcgggtgc gccgcgggtt cctggcgaca ctgggccagc 27600  
 cgtggctcac cggcggttta ggtcaggcgt gggcggttgc cagcatggcg ggtgcggtt 27660  
 tgcgtaggtc gggtaggcgc atccggcgcg ggagccggtc gagttcttcg ccgatggccg 27720  
 gtgctttggg gctgctcagg agccgaacac ctcccagccg caggtgccgg gctgaaccga 27780  
 gtggttctcg tcggctcgga tcacaacgtc tgccggaaca gctgcggcga ggtggtcgca 27840  
 gattcgaggc gggatcgcc tcggcgacct tgccgacgat cgcggctagg gcccgaggct 27900  
 tcgtcgacct ggttggcacc tagatcacga cggtaaaaac ttgccggcat cagagacgat 27960  
 cgaagtgatc ccgggtcacg tcggcttata ggtcgagtga gtcccggggc ctgccagcc 28020  
 aggtcttgcg tcgttgttcc gggctcagtt gcggattccg acgaacaggc ctccggcgtt 28080  
 cgggtgctcca ggaaggattt ccgcgcggat ccctgcgtct tcgagcgcg cggtgtactc 28140  
 gtcctcagtg aacagcgaga ggatttcgaa ctctgtgaag tcccggatcc cgggtgggtt 28200  
 ggcgactgtg tagcgacgg tcctccggct cgtacggccc tccaggaccg agtgcgatag 28260  
 ccggctgatc acccgctcgc cgtggtgcgc gacggctccg gtgacgaacc cgtcgatgaa 28320  
 cttgtcggga aaccaccagg gttcgatgac cgcgactcca ccaggggcca ggtgccgggc 28380  
 catgttccgc gtcacgcgtc gcaggtcgtc aacggtccgc atgtaagccg cggtaaagca 28440  
 caggcagggt atgacgtcga atggctcgc gaggtcgaaa tcgcggatgt caccgatgtg 28500  
 aatcggtacc tcagggactc gtctgatcgc gatctccgc atcgcatcgg acagttcaag 28560  
 cccgcgacc ttcgcgtatt cggcacggaa tcgctctagg tgcgccccg tcccacaggc 28620  
 gacgtcgagt agggactgtg ctccggcgag cctggtgcgt acgagctgga ctacttcccc 28680  
 ggctcggct gccagtcgc ggccacgcgc ggagtggatc gcgtcgtaga tgtcggcatg 28740  
 atctgggctg tataccgagg aggtttctgc gaatgtgtcg ctacgcgcgc acatcctcac 28800  
 ttcggagtgt gtgatctttg gctgatgtgg tgttcgacgg ccttctggaa ctgcgcagcc 28860  
 accgtgcga cctcggcgtc gtcaaggctt ggggtcagtg gtagcaggag tgttctgcgg 28920  
 caggcgtgct ccgcagaagg cagcttgcag tccgcgcggt agatggggac cttgtgcagg 28980  
 ggcggttagc ggtagctcgt gtagatgccg cgttccagca tttgctgcgc cacctggtcg 29040  
 cggatctccg gagccagctg gaccagtag aagtagtgtg acgagacgtg cccatccggt 29100  
 agcgtcggcg gtaggaggac acccggcaca tcggaaagca accggtcgta ctgcgtagcg 29160  
 atttctctac gcctgttgat gaattctggc agtttgcgca gctgcacgct gccaaagcgt 29220  
 gccgtcatgt cgttcccgat cagccgctgg ccgatgtctt cgacgcgaat atcccaccag 29280  
 cggttggaag acttggccga atcgaatccg ctcatctgct caagaccgtg gtaggcgagt 29340  
 cgtcttgcgc ggtgcgccag ctccggatcc gccgcgtaga acatgcccc atccccggtg 29400  
 accaggatct tcatcgcatc gaaactccac gtggccaggc caccaaaggc tccgcaagcg 29460  
 gtgcctgca cggacgatgc caccgcgcag gcggagtcc cgtatgagcat gaggcccttt 29520  
 tcacggcaga aatcggcgat cgcggtgact tctccggcg atcctccata gtggagcagc 29580  
 aatacggcct tggtcgccg cgtgatggcc ctgccacat catccagcgt ggggttcaac 29640  
 gtccgggggt cgacgtcga gaacaccggg cgggcaccgg aggatgcgat ggcgttggcc 29700  
 gccgccaga agcttatcga aggaagtacc acgtcgtcgc ctgggccgag gtcgagcacc 29760  
 tgcacggtaa ggaacagcgc ggcagtcgcc gagttgagga acacgacctg ttcgggatcc 29820  
 actcccagg ggtggggcga ttcggcctcg aacgtccggg tgcgcggccc gagcccgatc 29880  
 cagttggagg cgaacacctc ccgatcgcg tcgagttctt cgggtgccgag gatcggtgtg 29940  
 tgcaggttga tcacgttgct gaaatcctcc gagatgcgc catgctggat gctaggaact 30000  
 cttggccacg aattcagcga ttgattcgac gacgtagtcg atcatttggc ccgttatgcc 30060  
 tgggtagacg ccgaccaga aggttcggtc ggtgacgatg tcgctgttg tgagcgctc 30120  
 ggcatccgg taccgcacct gctcgaaggc cgggtgccgg gtgatgttac cggcgaacag 30180  
 cagtcgggtg ccgatgttg gggattccag gaagttcacc agggcgccac gggtaaccc 30240  
 gggtcccgca ctgatgtga tcgcaaaccc gaaccagctc gggtcgctgt gcggtgtggc 30300  
 taccggcagc agcaggccc gcaaccggga cagcccttcg cgcaacgctc gccagttacg 30360

gcggcgtgcc gacccgaatg cggaaatctt gctcaactgg ctcagcgcaa gtgcggcctg 30420  
 caggtcgggtg gtcttgaggt tgtaaccgac gtggggagaac gtgtacttgt ggtcgtagcc 30480  
 cgggtggaagg gtaccgaggt ggtagtcgaa cctcttgccg caggtgttgt ccacgccggg 30540  
 ctcgcaccag caatcccgtc cccagtcacg cagcgactcg atgatgcgag ccaattccag 30600  
 gctgccgggtc aacacgcagc caccctcgcc gctgggtgatg tgatgggcag gatagaagct 30660  
 gaccgttggtc aggtcgccga aggttcgggt cagccgtccc cggtaggtgg atcccaccgc 30720  
 atcacagttg tcttcgacga ggaacagctc gtgttctttt gcgatctccg cgatttcgtc 30780  
 agcggcgaaag gggttgcccc ggggtgtgcgc cagcatgatg gtcgcgctcc gttccgtgac 30840  
 ggcggccttg atgcgggtctg gcgttgcggt gtaggtgccc agttccacgt cgacgaatac 30900  
 cgggacgagt ccgttttgga ccgccggatt gatcgctcgtg gggaagccga ccgccgcagt 30960  
 gatcacttcg tcgccgggccc gcagtcgtgc ctcgccgagt ttgggggagg taagcgaact 31020  
 cagtgccagg agattggccg acgaaccgga gttgacgaga tgagccttgc ggaggccgaa 31080  
 gaagcggggcg aactcgctct cgaatcgccg tgcattcccc cccgcggcga tccggagctc 31140  
 cagcgcggct tccaccagtg ccacccggtc gtcctcgtcg agcacggcgc ccgatggccg 31200  
 gatcggcgctc gatccagcca cgaaggtcgg ggattcctgt tcgcgggtgg aatcgcgtaac 31260  
 ggatgccaat atccggtcct tggcatccgg caccatctca gtagcggtag cgcaagtgtc 31320  
 gtcacacgaa gtcactctgg cgcgcctttt cccagcgct ctggttttcc ggctctgcat 31380  
 gcaggcgacg atcagttctc gcgccttgcc ttcaggagat gagcgatgcc cgtggcgaat 31440  
 cgcgttatga cgtcccagcg ggacagtgtg ctgtctcgcc gccttacacc ttctgacct 31500  
 ggttcgatgc ggtgcgggac atcaggacag cggagcaagg agaagcgctc attgactcag 31560  
 aaatcctcga tctaccggc acaccgcact cggtagagcc caggctagcg ggaacgacct 31620  
 gtcgcgcctt gtcaagatcg ctaccatcac ctggaaggcc taagatttgg cttgcgaaag 31680  
 cggcgcttcc cgggggatat cagagatttc tgtgattctt ggcatgcttc ccgggtgttc 31740  
 aattgcgata ggagagttca tgcgtgtcct gttcaccccg ctgccggcga gttcgcactt 31800  
 cttcaacctg gtgccgttgg cgtgggcgtt gcgtgccgcg gggcacgagg tccgtgtcgc 31860  
 catctgcccg aatatggtgt cgatggtcac cggagcagga ctcaccgcgg ttcccgtcgg 31920  
 cgacgagctc gacctcatct ccttggcggc caagaacgaa ctcgttctcg gcagcgggggt 31980  
 ctcgttcgac gagaaggggc ggcattccga actcttcgac gagctgctgt caatcaactc 32040  
 cggcagagac acggacgccg tggagcaact ccacctgtg gatgaccgat cgctggacga 32100  
 tctcatgggg ttgcgcgaga aatggcagcc tgatctcgtt gtgtgggacg ctatggtgtg 32160  
 ttccggggcca gttgtggcgc gagcgctcgg cgcacgacac gtgcggatgc tcgtcgccct 32220  
 cgatgtgtcg ggggtggtgc ggtccggttt cctcgaatac caggaatcga agccgcctga 32280  
 gcagcgcgtc gacccgctcg ggacgtggtt gggagcgaag ctcgccaagt tcggagccac 32340  
 gttcgatgaa gagatcgtga cgggccaagc gaccatagat ccgattccat cctggatgcg 32400  
 cctgcctgtg gacttggact acatctcgat gcgtttcgtg ccgtacaacg gtccggcggt 32460  
 gttgccggag tggttgcgcg aacgaccgac gaagccgcgc gtctgcatca cgcgcgggct 32520  
 gaccaagcgg cggctgagca gggtgaccga acagtacggg gagcaaagt accaggaaca 32580  
 agcaatggtg gaaaggttgt tgcgcggcgc ggccaggctc gacgtcgagg tgatcgccac 32640  
 cttgtctgac gacgaagtac gggagatggg ggagttgccc tcgaacgtcc ggggtccacga 32700  
 atacgtaccg ctcaacgaac tgctggagtc gtgttcagt atcatccatc atggctcgac 32760  
 gacgacgcag gaaaccgcca cggtaacgg cgtaccgcag ttgattctcc ctgggacctt 32820  
 ctgggacgaa tctcgtaggg cggagctcct agccgatcgg ggagccggtc tggctctcga 32880  
 cccgcgacg tttaccgaag acgacgtgcg aggtcagctg gcccgcctgc tcgacgagcc 32940  
 gtcgttcgct gccaacgcgg cgtgatccg ccgtgaaatc gaggaaagtc ccagcccgcga 33000  
 cgacatcgtt ccacgtctgg aaaagctagt tgccgaacgt gagaaccgcc gcactgggca 33060  
 gtctgatggc catccgtgag caacgtgtgg ccggaacat ggacgccggg gtttggcagg 33120  
 tgttcacgc tgttgcgtcg actcggattc cgcctgacc gggacgatgc caggcgagtc 33180  
 ccgaagtcag attcttgtcc agaatcgctc aatgggggtg tgatctcccc agaggtttgc 33240



gctccaaccg atttccgacg aggatcgtgg cgcccgtga gcaacgacta ccgtgcggtc 33300  
gagacatacc gctgtgcgcc aggagcgaag gtgggttgcc cgatcaccgt gctgggtgta 33360  
gatgccgagc cgaaggtcac cttggatgag gcggaagcct ggcgagagca caccgaggcc 33420  
gtggccgacg tccgtgtctt ctccggcggg cattttcttca tgaccgaacg ccaggacgag 33480  
gtgctcgcgg tccctacggg cggatcgctt cgatgatcct cgccaggccg ctggaccaga 33540  
ccgcgacgcc cctgggagcc ggcgtgcaca tcgtcacggc agtgagggat tgggcatgag 33600  
cagttctgtc gaagctgagg caagtgtgtc tgcgccgctc ggcagcaaca acacgcggcg 33660  
gttcgtcgac tctgcgtga gcgcttgcaa tggcatgatt ccgaccacgg agttccactg 33720  
ctggctcgcc gatcggctgg gcgagaacag cttcgagacc aatcgcatcc cgttcgaccg 33780  
cctgtcgaaa tggaaattcg atgccagcac ggagaacctg gttcatgccg acggtagggt 33840  
cttcacggta gaaggcctgc aggtcgagac caactatggc gcggcaccca gctggcacca 33900  
gccgatcatc aaccaggctg aagtaggtat cctcggcatt ctctcaagg agatcgacgg 33960  
cgtgctgcac tgcctcatgt cagcaaagat ggaaccgggc aacgtcaacg tctgcagct 34020  
ctcgccgacg gttcaggcaa ctcgagcaa ctacacgcag gcacaccgtg gcagcgttcc 34080  
gccctatgtg gactacttcc tcggggcggg ccgcggccgc gtgctggtag acgtgctcca 34140  
gtctgaacag gggctcctggt tctaccgga gcgcaaccgg aacatggtg tggaagtcca 34200  
ggaggaagtg ccagtcctgc cagacttctg ctggttgacg ctcgccagg tctggtctct 34260  
ccttcgtcag gacaacatcg tcaacatgga caccggacg gtgctgtctt gcaccccggt 34320  
ccacgattcc gccaccggac ccgaactagc gcctcggag gagcccttcc gacaggcggt 34380  
ggccaggteg ctctcgacg gcacgatc gtcgagtatc tccgaggcgg tcggttggtt 34440  
cgaggaagcc aaggcccgt accgcttgcg ggcaacgcgc gttccgctga gcagggtcga 34500  
caagtggat cgcaccgata ccgagatcgc ccaccaggac ggcaagtact tcgcggtgat 34560  
cgcggtgtcg gtgtccgca ccaatcgtga ggtcgccagc tggacgcagc cgatgatcga 34620  
accgcgagaa caagggtgaga tcgactgtt ggtcaagcgg atcgcgcgag tgctgcacgg 34680  
tttgggtccac gctcgggtgg aggtgggtta taagtggact gcggaaatcg ctcccacggt 34740  
ccagtgcagt gtggccaact accaaagcac ccgctcgaa gactggccgc cgttcttgga 34800  
cgacgtgtc accgccgatc ccgaaaccgt gcggtacgaa tcgatcctgt ccgaagaagg 34860  
cggctcggttc taccaggcgc agaacaggta ccggtatc gaggtgcatg aggacttcgc 34920  
ggcacgacct ccagcgact tccggtggat gactttggga cagttggcg agctgctccg 34980  
gagcaccac ttcttgaaca tccaggcgcg cagcttggtc gcctccctgc atagcttggtg 35040  
ggcgttgggg cgatgaccag ctcgatgcga aagccggtgc gcacgggtgt gctcgggtgc 35100  
gcttccttcg cgtggcgacg gatgctgcc gcgatgtgcg acgtggccga aacagagggtg 35160  
gtggcgggtg cgagccgtga tccggcgaaa gccgaacggt tcgcagcgcg attcgaatgc 35220  
gaggcgggtgc tgggttaccg gcggctcctg gagcgccgg acatcgatgc cgtctacgtg 35280  
ccgttgccgc ctggcatgca tgcagagtgg atcggaagg cgcttgaggc agacaaacac 35340  
gtgcttgcg agaaaccgt gacgacgac gcgtccgaca ccgctcgctt ggtcgggctg 35400  
gccaggagga agaacctgt gctgcgggag aattacctgt tctccacca cggccggcac 35460  
gacgtggtcc gcgacctgt gcaatccgg gagatcgggt agctccggga gttcaccgcc 35520  
gtgttcggaa ttccgccgt tccgacacg gacatccgt atcgaccga actcgggtggc 35580  
ggagcgttgc tggacatcgg tgtctatccc gcccggtcgc ctcggcactt tctcctcggt 35640  
ccgctcacgg ttctcggcgc aagctcgac gagggccagg agtcgggct cgacttctcg 35700  
ggcagcgtgc tgcctcaatc ggaagggtgg accgttgccc acctcgata cggtttcgtg 35760  
caccactacc gcagcgcgta cgagctgtgg gggagtcgtg ggcgaatcgt cgtcgaccgg 35820  
gcgttcacgc cgcgcgccga gtggcaggcc gtgatccgaa tcgagcgga gggcgttgtc 35880  
gacgagttgt ccttgccagc ggaagatcag gttcgcaagg cggtcaccgc cttcgacgc 35940  
gacatcagag cagggacagg cgtggacgac cctgcgggtg ccggagattc gggcgaatcg 36000  
atgatccagc aggcgcgct ggtggaggcg atcggtcagg ccgctcgggt cgggtccaca 36060  
tagccgcccg gcacccgcg gtagtagttc gcctcgaagc ctgacgggc atccggaagc 36120



cagcggggaa gccgctggag aggtcaccg ccatccgctc acctggcatc tcgcgagaccg 36180  
ctgatcgcgg acggctcgga gaagtgtctg tcgaaccacg agacgaccac tcgcgagctg 36240  
gccagggcgg cgggaaagtg agccaatccg gagagcggat gccaccgcac tggcgtaccc 36300  
gccgcgcggg agctgtcccg gagtgcgtcg ccgaatgcga acggaacgat ctgctcgtcc 36360  
gtgctgtggg agacgagcgt ggggaccacc gggccaccgt tcctacctgc gacgctttcg 36420  
gccagtcgtg cgcgccatcg aggttgctcg aaaaggccgg aagtgtcgag gaagtcgctc 36480  
agctcgcggc cgaggaagcg ggtgacgagc tccggtgcac cgagctcgcg cacttgatca 36540  
acggcggtag gaccgccttc ggtgagaagc tcgtcgaatg gcagatcggg gtaggcagcg 36600  
gcatgcccgga ccaggccggc cagcaccggc ccggtgaaca ccccgctatt tcgggtggatg 36660  
atgtccagca gatcgatcgg caccgcacct gcggccgcag cgcggattcg cagttcaggt 36720  
gcgtaggtgg ggtgcagttc gccggcgaag gccgacgctt gccaccctg cgcatagcc 36780  
cagatgccga ccgggcagtc ggtcgtcagg ccggagcccg gtagccgttg cgcagcgcgg 36840  
gcggcatcga gcatggcgtg tccctgcgcc ctgccgacgg ttaggtgtg ggttccagga 36900  
gtaccgaggc cttcgtagtc ggtgatgacc acggcccacc cgcggtcgag ggccacggcg 36960  
atcagctcgg tctccggctc ggttccggtt cgaagcaggt acgacggggc aacttggtta 37020  
ccgaggccgt ggggtgccac tgcgaaagtg atgatggggc gatcttcgcg cggccaacggg 37080  
atgttcggca ccagaacggt gccggagacg gcgttcggca tgccaagggc ggagttggac 37140  
cggtagagga tttgccaggc cttggctgcg acgggttcgc ccgtgccgcg cagtgcgcag 37200  
acgggcccgg ccctgaggag cgtgcccggg acaccggcg gtagcggcgt cggcggtcgg 37260  
tagaagggat catccgcggg tgcgccgaga tcgtcgcgga ccaggctggc gtgctcggag 37320  
gccatcagga ctgcttcttt cgagcctgca ggagcatgaa acccatgctt tctcgtttc 37380  
tggcgtaatc cggtatgttc cgttattccg caaccgcggc gatcagctgt gctggtccc 37440  
gtccgtgctt cgcgcgatg tctcccaagt agcgttgctg gtaggtgccg acagccgcag 37500  
gctcgacgcc ggcgagctca tcgagtttcc ggagcaactc gtcgacgtac caggagacca 37560  
tgcacctggt ctgtgccgtg aggtcgggtga cttcgagaat ctogaaccgc gcttcgctga 37620  
ccagcgcgtg gaagctgttc aaggtatggg cgtcgtgcc cgtccaaacc gccgcgtact 37680  
cttccgggag tcgaaccgga gtgatgatgt ctccgaggac gaaccggccg ccgggttcca 37740  
ggattcgggtg gacctcgcgg atcgcggcgg cctggtccac gatctgcacg acggactgca 37800  
tcgcccattg gccctgaaag aaaccgtccg ggtagggcag ctgggcgcgg tcgactagat 37860  
cgaactcaag actgccggcc agtcgggtct cgttggcgag cctggtggcg gcggcgagat 37920  
gctgggcgtt cacggtgatt ccggtgactc gaacgccgct ggcgcatgcc gcacggacta 37980  
cgggctgccc attgccgcag ccaggtcga acaggtgcgc tccgggacgg agcgcggcct 38040  
gtcgatgaa caggtcgttc agttggtcgg cagcatccga ccacggtgtg gcaccggcat 38100  
cctcccata cccgcccgcc cagtaaccgt ggtgcagggg acgcccgctc gccaacgcat 38160  
cgaagatgga ctccacctga tccgcggtcg gaaatgcctg tgtgttcgcc cctctgctgt 38220  
tcactcgtcc tccgcgctgt tcacgtcggc caggtgcaat atgtcgtcca gactccttgg 38280  
cacccaagca ggaacgcgc cttcggcgtt gacgccttcc tccaggaacg cgatgttgtg 38340  
gtaggtgtgg aggcgcacca aattgcgttc caggtagctc ggctcgtacg agcccgcatg 38400  
cggtgctcc tcgtgctgaa cgccttccaa caggttcttg agcaggtga ccgtggtgcc 38460  
gggtgcggcc gggcactgcg cctgcccgcg gaatccggga gcataggtcg tccacagatc 38520  
ctcgatcacg tatacgcac cgctgcgcaa ccgggggaac agcgtttcca gggatgtgcg 38580  
cacgtgtccg ttgatgtggc tgccatcgtc gatgatgatg tcgaacggtc cgtacttgtc 38640  
gtcaacggcg gccagctcct cgggcttgct ctggtcggcg cggacggtgc agagcctctg 38700  
ctggtcgagg aaggacttgt cgaaaacgtc catcccgaac acgaggccgc ggtggaagta 38760  
gcgttccac atcttcaggg attcgcgcgc gccaccgtcg aagttgtagc caccgacacc 38820  
gatctccagg atgcgcaccg ggcgatcacg gaactcgcgc aggtgtcgtc cgtatagcgg 38880  
ggtgaaccag tgcaggccgc ccacttgct cgtgcggtag tgggaggcga gcaagttgag 38940  
gtcgggacgt ccggtgccgc agccggcgac cactgcggag atggcctgga agccatcgga 39000

cagttccgac ggaccgggta tcgaaccgga tgtggtggtt cggaggaagt tgggtgctccg 39060  
 ggcgcgcgacg gccctgggag ctccctgggccc gaacaactcg gcgatgagat cgggtgagctc 39120  
 gtaaccgatac cgcagcggga cgtctccgac cggtcggttc tcggccttga tcagctcacc 39180  
 ggactgtagc gtcaggacga agtcaacggg ctccgctcgg tgggtgatct ggaccgcgac 39240  
 ctccggtccgt tcgatgtcgg gggccgggtc cgcgcggaag aggatctcgt cgatcagcac 39300  
 ggggtgcgac ctggcgagtc cgagttcggg ggtcaggteg gccaggctcg ccgcactgga 39360  
 tccggcggcg aggatgatgc gttccacggg ttccgatctcg tgcgttggtg acatcgtgat 39420  
 gagctcctca tgggtgaccg ggtgaaagcc gtgccggcgg tttgatcgac aggccgtgct 39480  
 ggaagatggt ctgocgatcc caccgcgctt tggcccgtcg cagccgcggg tagttgtctt 39540  
 tgtagtacag gtcgtgccag gcaacaccgg aggtgttcca caatggatcg gccaaagtccg 39600  
 tgtccgggta gttgatgtag gagccgtcga caggggtacc tggcaccgga actccgcggg 39660  
 tttcggcgta catctcggg tagaaaccgc gaatccaggt cagatgccgc tcgtcctcgg 39720  
 cgggctccga ccagttcgtg acgaacagcg ctttgagaac cgagtcgcgc tgagcgagtg 39780  
 cggtgggcca cggagccacg gcattcgcca taccgccgta accgagcagc aacagcgccg 39840  
 ccgcagggtt gtcgtatccg tagacggta gccgcgggta aaccgtggct agttgagctt 39900  
 cggacagccc ggtgcgcaag taggocggctt tgacctggt ccgttgcatg cccggttcgc 39960  
 tgccttcggc gatcgcctcg gccacctggg tcgatcgcaa ccacggcagg gtttcccga -40020  
 gcccttcggc tggagtcacg ccgacctggg cgttgatcgc cgacagggtg tcggccaggg 40080  
 tgcgttccgc gttcggatcc gtgccgtcca ggtgaacggt cagcgtgacg tagccagctt 40140  
 gccgggtgtg gcagacgagc gtgctgaaca acccgagttg cgtggattca ggcgcgctgt 40200  
 gctgctcgta ccaattgccg aagttctgta ggagcacggc gaatgactgc tctgtcagtt 40260  
 cgtgccacgg ccagtggaaac gatcggagca gcaactgtcgc gggcggccgt ggcaggagct 40320  
 ctgocggcgc ggtgctgacc acgtccggcg ttcggagcca aaacctggg acgatccga 40380  
 agttgcgcgc accgccaccg gtgtgcgcgc accacaagtc gtgaccggcg cccgtggagt 40440  
 tccggtcggc ctcgacgatg tgcaactcac cggcctggtc gaccacgacg acctcgacgc 40500  
 cttgaaggta gtcgacgacc gaaccgaatc ggcgcgacag cgggcccgtat ccccccgcga 40560  
 ggatgtgccc gcctgcgccc accccgggac atgcgcgggt cgggatcgtc acgccccagt 40620  
 tcttgaacag ggttcggtag acctgcccga gggcggcgcc cgcctcgatc gcgaatgccc 40680  
 cgcgcgtgct gtcgtagtac acgcggttga gtcggagag gtcgacgagc actcggatcg 40740  
 ccgggtccgc aacgagattc tcgaagcagt gcccgccgct gcggaccctt accgcctgc 40800  
 cggtgccgac ggcgtcggcg acggcgtgca cgacgtcttc ggcgagctg gcgatgtgga 40860  
 tgcgttcggg ttttccgggtg aaacgggggt tgtgcccgac gacgaggtcc ggataacgag 40920  
 gatcgtcggg ctcgacgggtg atctctgttc ctgggggttc acgattcatg ggtgccgggt 40980  
 catggaattc gggcaccgcc cctccttttc tgactggtcc actttgttcg ccgcagccg 41040  
 agatcatcta cgcgtccggg tgattatctg tgtgtttcag ctcatcagtg aaaccggtc 41100  
 gcctccgccg gctctacttt gtggatcgat atcgcggtgc gcatggtgcc gtatgcgctg 41160  
 gaaccgaaaa ggtgatgact ttccatgagt gagatcgag ttgccccctg gtcggtggtg 41220  
 gagcgtttgc tgctcgcggc ggggtgcggc ccggcgaagc tccaggaagc agtgagggtg 41280  
 gccggactgg acgcggtggc cgacgccatc gtgcagcaac tcgtcgtagc ctgcgatccg 41340  
 ctgtcgttgg acgagtcggt gcgaatcggc ctggagatca cttctggcgc tcagctggtc 41400  
 cggagaaccg ttgagctcga tcacgcaggc ctgcggctcg cggcggtcgc cgaagcagct 41460  
 gctgttctcc ggttcgacgc ggtggatctg ctggaagggc tcttcggccc ggttgacggc 41520  
 aggcggcaca acagccgtga agtccgctgg tcggacagca tgacgcagtt ctgcgccgac 41580  
 cagggcctcg ccggcgcgca gcgcctgctg gcgttcggga acagggtgtc caccgcgggtg 41640  
 cagcccggtc tggccgcagc cgccaccagg cgcgcggacc tcggtgcgct ggcagtcgc 41700  
 tacggatccg acaaattggg ggacctgcac tggtagaccg aacactacga gcaccacttc 41760  
 tcccgaattc aggatgcccc ggtgcgagtg ttggaaatag gaatcgggtg ttatcacgca 41820  
 cccgaactcg gtggtgcttc gctgcgcatg tggcagcggt acttcggcg aggtctcgtt 41880

tacgggctgg acatttttoga gaaagccggg aacgaagggc accgagtgcg aaagctgcga 41940  
 ggtgaccaga gcgatgcgga attcctggaa gacatggcgg ggaagatcgg gccgttcgac 42000  
 attgtcatcg acgacggcag ccatgtcaac gaccacgtca agaaatcctt ccaatccctg 42060  
 tttccgcacg tccgcccagg tggtttgtac gtcatcgagg atctccagac ggcgtactgg 42120  
 cccggctacg gcggtcgcga tggggaaccc ggggcccagc gcacctcgat cgacatgctc 42180  
 aaagaactga tcgacggcct gcattatcag gagcgcggaat cgcggtgcgg gaccgagccc 42240  
 tcctacacgg aacggaacgt ggcggccctg cacttctacc acaacctggt attcgtggag 42300  
 aaagggctca acgctgagcc tgccgcgcgg gggttcgtgc cccggcaagc gtcggcgctc 42360  
 gaggacggct gagccgttca ccagctgcgg cgccagtagg cggccgtgcc gtcgatgtcg 42420  
 tggatgggtt ccgtgatccc gagttccgcg cggaacccct tcaccgcgtc ctggcaggac 42480  
 ggcagaaaat agtcgtcgat gatgacgaat ccgcccggcg agagcttcgg gtaacaggttc 42540  
 cgcaatgagt ccattgtgga ttcgtagagg tcgccgtcga gtcgtagcac ggcgagttcc 42600  
 tggatggggg cgggtgggcaa ggtgtcccgg aaccagccgg ggaggaaacct gacctgttcg 42660  
 tcgagcagcc cgtagcgggc gaagtctcgc cggacgggtc caagcgatac gccaaagcacg 42720  
 tcgttgact cgtgcagcgc catagcctgg tccgcttggg ggtcttgccg agagctttcc 42780  
 ggcattccct ggaaggaatc cactaccag acggtacgtc cggtatctcc gaatgcctgg 42840  
 agaaccgcgc gcatgaagat gcatgcgcgg ccccgccaga caccggtctc ggcgaaatcc 42900  
 ccgggaacac cgtctgcgag cacggcttcc acgcagtgtt ggaggttgtc cagccgctcc 42960  
 agaccgatca tcgtgtgcgc gacagttggc cagtcctgtc ctttggcccg agcggcctgc 43020  
 ctgtagtcgg tgttgtcctg ccaggcgttc ggatgcggcc gatcactgta aatcgtgttg 43080  
 gtgagtacct tcttgagcag gtccaggtag agcgcgttct gggagggcat cggttctccg 43140  
 gatccagctg ttctcgggtg actagttcat caggcacgga tggccgcagt gttctccagc 43200  
 gtccgcacca gcgcggcggg atggggcatg gccgtgatct cgtcgtgag tttgattgcc 43260  
 gcagcagcga agccggtgtc gccgagcacc gttgcgattg agtcggtgaa ctgttcgtgg 43320  
 tcggactggg cctgctcatc cggcaagcag atgcccgccc cggcagcggc gaggttgccg 43380  
 gcgtagtcga actggtcgaa gtactgggga agcacagagt gcgggatgcc gagtgcgggtc 43440  
 gcggtgaatg ccgttcccga gccgcccgcg cagatgacca gctcgcaggt acgcaggaac 43500  
 aggttgagcg ggaccgatcc ggcgatccgg gcgttgtcgg gtaggtcggg gagaagtgcc 43560  
 cgggtgctcg ggggaacggc gatcacggcc tcgacgcggg gcaactcggg ggcagccgct 43620  
 actgcgcgca gcagcggagc cggcccgggt gcgttcagca ccatgcggcc catgcagatg 43680  
 cagaccgcgc gtgctgaggt gcgcgcgcgg ccccatgccc ggaatgcgcc gcttccgttg 43740  
 tacggcacgt actggaccgg tgccgcttgc ggcgcgtcgc ttgcttgag gctcggcgga 43800  
 cagggatcga ggatgagctc gggagtgggc aggcgggtca gtccgtgggt ccggcacacc 43860  
 ggggtcaagca actcgtgggc tcgatcgctg aaggggcctg cgggtggggtc gactccccag 43920  
 cgggtgcagca cgaccggcag gtcgagcaat ccgccgagca cccggccgat cagcgcgcag 43980  
 acgtcgacca acagcactga cggtcgccag gcctcggcca gtcgaaggta ttcggggagc 44040  
 tgatcgagcg agctttgcgc gacattggac gcggtctgct cccacagttg ccggcctgcc 44100  
 tcggtgtcgc gctgaccgaa cgccggattg ggaaagcgca gctgcgtggt tccaccgta 44160  
 tcgcccgtcc tgtcgttccc gcggatcccg gccgtggtga gacctgcacc atgcgcggtc 44220  
 gcctgcagct ctggtgggtg ggcgatcagg acctcgtgcc cggatgcttg cagcgcgccg 44280  
 cacagcggca ccattgccat gagatgcgtc ggatagggca agggaaacgac gagtacgcgc 44340  
 atacttcgga cccagctctc tttccccga ttagcgcagc agccctact cccattggcc 44400  
 aggatttgga aaatgcgctg cgtatgtcga tcgccgttga cgtccaacgg acttccggcg 44460  
 gcaacaatag tgtgtcacgg caggaatgtc acgcgacat cgaagatctt tgggtcgccg 44520  
 cacctggttt cacgcgaacg agtgaaatgc gcgagctccg ctcgatcggg gtgggcccga 44580  
 cctgtacggg gatcaccgtt ggttctgcgg ggattcatgg ggaagatttg cgtggtgt 44640  
 ttgcctcctg gccggatagt tatagtcggg accgccgat gcggcggtaa ccgcgaatta 44700  
 actgacggct agtttgccgt ctttctctc tgtgtgttct ctgctcggtt ccagaaaatt 44760

acgagaaggt gaacgttgca gagatcaggc ataccggtgt tgccaggtgg cgcaccaaca 44820  
 tcgcagcagg ttgggcagat gtatgacctg gtcacgccgt tgctgaactc ggctcggggc 44880  
 ggccccctgcg ccatccacca cggctactgg gagaacgacg ggcgggcttc ctggcagcag 44940  
 gccgcccagacc ggctcaccga ccttgctgcc gaacggaccg tgctcgatgg cggcgttcga 45000  
 ctgctcgatg tgggggtgcg taccggacaa ccagcgctgc gcgtcgcgcg cgacaacgcg 45060  
 atccagatca ccggcatcac cgtcagccag gtgcaagtgg ccatcgccgc tgattgcgca 45120  
 cgcgaacgcg gactaagcca ccgggtggac ttctcgtgcg tcgatgccat gtccctgccg 45180  
 taccgggaca atgctttcga cgccgcctgg gccatgcagt cgctgttgga gatgtccgaa 45240  
 ccggaccgtg ccatccggga aatccttcga gtactcaaac ccgggtggcat cctcggcgtc 45300  
 accgaggtcg tcaaacgaga agcggggcggc gggatgccgg tgtccgggga caggtggccg 45360  
 accggccttc ggatctgcct ggctgagcaa cttctggaat cgctgcgtgc agcgggggttc 45420  
 gagatcctcg attgggagga cgtgtcgtcg aggaccgggt acttcatgcc gcagttcgcc 45480  
 gaagagctcg ctgcgcacca gcacgggatc gcggacaggt acgggcccgc tgtcgccggc 45540  
 tgggccgcgg cggctctgca ttatgagaaa tatgccacg acatgggcta tgcgattctg 45600  
 acggcgcgga agccggtcgg ctgagggcgc gccgaattc gatgacgttc atgcgccgtg 45660  
 tcggggaatc gccggtggcg gcgccagcag aggtgaact tactggtggt gtgtccagga 45720  
 atcggagggg cagtaccgaa tgagcgaagc cgggaacctg atagccgtca tcggactgtc 45780  
 ctgccccta cccagggcgc ctgaccggc ttccttctgg cggttgctgc gcaccggaac 45840  
 ggacgccatc accacgggtcc cggaaggcg gtggggcgac ccgttgccctg gtcgggatgc 45900  
 gcccgaaggc ccggaatggg gtggtttcct ggctgatgtc gactgcttcg atcccagatt 45960  
 cttcgggatc tcgccgcgag aagcggcagc cgtggatccc cagcagaggc tggctctgga 46020  
 gctcgccctg gaggcactcg aagacgccgg tatccccgcc ggcgagctgc gcggtactgc 46080  
 cgccggtgtg ttcattgggg cgatctctga cgactacgcc gccctgctgc gcgagagccc 46140  
 gccggaagtg gctgcgcagt accgcctcac cggcacccat cgaagcctga tcgccaaccg 46200  
 cgtgtcctat gtgctcgcc tcgcggggcc aagcctgacg gtggattcag gtcagtcctc 46260  
 gtccttggtc ggctgcatc tcgccagcga gaggctgcga cggggtgagt gcacgatcgc 46320  
 actcgccggc ggctgaacc tcaacctggc cgccgagagc aacagcgctc tgatggactt 46380  
 cggcgcgctc tccccggagc gtcgctgctt caccctcgat gtgcgggcga acggttacgt 46440  
 ccgtggtgag ggccggcgcc ttgtcgtgct gaagaaggcc gatcaggcgc acgccgatgg 46500  
 cgaccggatc tactgcctca tccgcggcag cgcggtcaac aacgatgggg gcggtgccgg 46560  
 gctcaccgtt ccggcgggcg acgccaggc ggagctgctg cgccaggcat accggaacgc 46620  
 gggcgctgac ccggccggcg tgcaagtatg cgagctccac ggcagcgca ccagggtcgg 46680  
 ggatcccgtc gaagcagcag ccctcgagc tgtcctgggg gcggcgagac ggcccggcga 46740  
 cgagctgctg gtggggtcgg cgaagaccaa cgtcggccat ctggaagcag cggcgggcgt 46800  
 caccgggttg ctgaagacc cactcagcat ctggcaccgc gaactgccgc cgagtcttca 46860  
 tttaccgcc cccaaccgg aaatcccgt ggacgaattg aacctacgcg tccagcgtga 46920  
 tctgcggccg tggccggaga gcgaggggcc gctgctggcc ggcgtcagcg ccttcggaat 46980  
 gggaggcacg aactgccacc tgggtgcttc cggcacgtcc cgggtggagc gacggcgcag 47040  
 tggaccgct gaggcgacca tgccgtgggt cttgtcgcc agaacaccgg tcgcattgct 47100  
 tgccgaggcg gcgcgcttgc acacgcacct caatacgcc ggtcaaagtc cgttggaagt 47160  
 cgcctactca ctggcgacca ctgatccgc gctgcgcac cgggcgcgc tggtcgcgga 47220  
 cgacgaaccg aaactgctcg ccgggttgaa ggccctcgt gacggcgacg acgcgccac 47280  
 gctgtgccac ggcgcgactt ccggcgagcg ggcagcggtc ttcgtcttc ccggacaggg 47340  
 cagccagtgg atcgggatgg gtaggcagct gctcgaaacc tccgaggttt tcggggcgtc 47400  
 gatgtcggac tgcgccagc cattggcgcc gcacctgat tggccctgc tggatgtgct 47460  
 gcgcaacgcg gccggcgctc gcacacctga ccacgacgat gtcgtccagc ccgcgctgtt 47520  
 cgccatcatg gtctcgctcg cggagctctg gcgttcgtgg ggctgctgc cgggtggcgt 47580  
 cgtcgggcac tcgcaggggg agatcgcgcc ggccctgcgt gccggggccc tgtccgtccg 47640

cgatgccgcc aggggtggtgg cgggtgcgcag caggetttctg acggcgctgg ccggcagtg 47700  
 cgcatggcc tcgttgccagc atccccccga agaggtgcgg caaatcctgt tgccttgccg 47760  
 cgatcggatc ggcgtggcgg ggggtgaacgg accgtcgtcg accctgggtgt caggggaccg 47820  
 ggaggcgatg gcggaactgc tggccgagtg cgcagaccga gagctccgga tgcgccggat 47880  
 tcccgttgaa tacgcctccc attcgctca catcgaggtt gtcgggatg agctgctggg 47940  
 gctgttgccg ccggtcgaac ccaggacggg aagcatcccc atctattcga cgacgaccg 48000  
 ggacctgctg gaccggccga tggacgccga ctactggtac cgcaaccttc gtcaaccgg 48060  
 gctgttcgaa gcggccgtcg aggcctgtt gaagcggggg tacgacgcat tcatcgagat 48120  
 cagcccacac ccgggtgctga ctgcgaacat ccaggaaacc gccgtgcgag cagggcgagg 48180  
 ggtagtgccg ctccgggacac tccgcccgcg cgaaggtggc atgcccagg cgctgacgtc 48240  
 gctggccaga gcacacgtac acggagtggc cgcggactgg cacgcggtct tcgccgttac 48300  
 cggggcgag cgggtcgacc tgcgcagta cgcctttcag cgacagcgt actggctgga 48360  
 cgcaagctt ccgcagctcg ccatgcccga gagcgacgtg tcgacggcgt tgcgggaaaa 48420  
 gctgcggtct tcgccgaggg cggacgtgga ctgcagacc ctcacgatga tccgggcaca 48480  
 ggagccgtg gtccctggcc actccgatcc gaaagaggtg gaccggatc ggacgttcaa 48540  
 ggacctggg ttcgattcct cgatggtggt cgagctgtgc gaccgcctaa acgccgccac 48600  
 aggtctgca ctgcaccga gcgtcgtttt cgactgtcct acgccggaca agctcgccc 48660  
 ccaggtacgg acgttggtgt tgggcgagcc ggctcccatg acgtcacacc ggccggactc 48720  
 cgatgcggac gagcctatcg ccgtgatcgg gatgggctgt cggtttcgg gtggggtgtc 48780  
 ctgcccagag gagttgtggc agttggtcgc cgtgggcgg gagctcgtgt ccgagttccc 48840  
 ggctgaccga ggttgggacc tggagcgtgc ggggacatcg cacgtgcgcg ccggcggtt 48900  
 cttgcatggc gcccggatt ttgacccgg gttcttcgg atttcgcgc gcgaggcgtt 48960  
 ggcatggat ccacagcagc ggttgctgct ggaaatcgcc tgggaagcag tcgaacgagg 49020  
 cgggatcaac ccgcagcacc tgcacggaag tcaaaccggg gtcttcgtcg gcgcgacctc 49080  
 cctggactac gggccacgcc tgcacgaag gtccgaggag gcggccgggt acgtgctcac 49140  
 cggcagcacc acgagtgtgg cgtcgggtcg ggttgcgat tcgttcgggt tcgagggccc 49200  
 tgcggtgacg gtggatacgg cgtgttcgtc gtcgttggtg gccctgcatt tggcgtgtca 49260  
 gtcgttgctg tcgggtgagt gtgatctggc gttggccgg ggtgtgaccg tgatggccac 49320  
 gccggggatg ttcgtggagt tttcgcggca gctgggtttg gcgccggatg ggcggtgcaa 49380  
 gtcgttcgcg gaggcgcgc acggcaccgg ctggtccgag ggtgctggcc tggttctact 49440  
 ggagcgggtg tcggatgccc ggccgaatgg gcatgaggtg ctggcgggtt ttcgtggtag 49500  
 tgcggtgaat caggacgggt cgtcgaatgg tttgaccgcg ccgaatggtt cgtcgagca 49560  
 gcgggtgatt gcccaggcat tggcgagtgc ggggttgctg gtgtccgat tggatgctgt 49620  
 ggaggcgcat gggacgggca cgcggcttg tgatccgatc gaggcgcagg cgtgatcgc 49680  
 cacctacggc cagggccggc ttccggaac gccattgtgg ttgggctcga tgaagtcgaa 49740  
 catcggtcac gcgcaggcag ctgcggggat agccggcgtc atgaagatgg tgatggcgat 49800  
 gcggcacggg cagctaccgc gcacgttgca cgtggatgag ccgacttctg gggtgattg 49860  
 gtcggcggg acggttcaac tccttacgga gaacacgcc tggcccgga gtggtcgtgt 49920  
 tcgtcgggtg ggggtgtcgt cgttcgggat cagtgtact aacgcgcacg tcatcctcga 49980  
 acagcccccg ggagtgcga 50000

<210> 2

<211> 50000

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 2

gtcagtctgc ggggccgggt tcgggctctg tcgtggatgt tccgggtggtg ccgtggatgg 60  
 tgtcgggcaa aacacccgaa gcgctatccg cgcaggcaac ggcgttgatg acctatctgg 120  
 acgagcgacc tgatgtctcc tcgctggatg ttgggtactc gctggcggtg acacgggtcgg 180  
 cgctggatga gcgagcgggt gtgctggggt cggaccgtga aacgttggtg tgcggtgtga 240  
 aagcgctgtc tgccgggtcat gaggtctctg ggttggtgac cggatctgtg ggggctgggg 300  
 gccgcacggg gtttgtgttt tccggtcagg gtggtcagtg gctggggatg ggccggggggc 360  
 tttaccgggc ttttccgggtg ttcgctgctg cctttgacga agcttgtgcc gagctggatg 420  
 cgcactctggg ccaggaaatc ggggttcggg aggtggtgtc cggttcggat gcgcagttgc 480  
 tggatcggac gttgtgggcg cagtcgggtt tgttcgcgtt gcaggtgggc ttgctgaagt 540  
 tgctggatc gtgggggggt cggccgagtg tgggtgtggg gcattcgggt ggcgagttgg 600  
 cggcgcggtt cgcggcgggt gtggtgtcgt tgtcgggtgc ggctcgggtg gtggcgggtc 660  
 gtgcccgggt gatgcaggcg ttgccgtctg gcggtgggat gctggcggtg cctgctggtg 720  
 aggagctgtt gtggtcgttg ttggccgac aggtgatcg tgtggggatc gccgcgggtc 780  
 acgctgcggg tgcggtggtg ctctctggtg atcgggatgt gctcgatgac cttgccggtc 840  
 ggctggacgg gcaagggatc cggtcgaggt ggttgcgggt gtcgcacgag tttcattcgt 900  
 atcggatgga tccgatgctg gcggagttcg ccgaattggc acgaaccgtg gattaccggc 960  
 gttgtgaagt gccgatcgtg tcgacctga ccggagacct cgatgacgct ggcaggatga -1020  
 gcggggccga ctactgggtg cgtcaggtgc gagagccggt ccgcttcgcc gacggtgtcc 1080  
 aggcgctggt cgagcacgat gtggccaccg ttgtcagact cggtcggac gggcggtgtg 1140  
 cggcgctgat ccaggaaatg tgcgccgat ccgatcacgc cggcgcggtg agcgcggtcc 1200  
 cggcgatgag caggaaccag gacgaggcgc agaaggtgat gacggccctg gcacacgtcc 1260  
 acgtacgtgg tgggtgcggtg gactggcggt cgttcttcgc cggtaacaagg gcgaagcaaa 1320  
 tcgagctgcc cacctacgcc ttccaacgac agcggtagct gctgaacgag ctgctgtaat 1380  
 cttccgccgg cgacatgggc aggcgtgtcg aagcgaagtt ctggggcgcc gtcgagcacg 1440  
 aagatgtgga atcgcttgca cgcgtattgg gcattgtgga cgacggcgct gctgtggatt 1500  
 cctgagaag cgccttccg gtggtggccg gttggcagcg aaccgcacc accgagtcca 1560  
 ttatggatca gcggtgttac cgaattggct ggcggcaggt agccggactc ccgccgatgg 1620  
 gaactgtttt cggtagcttg ctggtcttcg cgcctcatgg ctggtccagc gaaccggagg 1680  
 tgggtgactg cgttacggca ctgcgggcaac gtggtgcctc ggtggtgttg gtggaagctg 1740  
 atcccgaacc gacctcttc ggcgaccggg tacgaacct gtgttcgggc cttccggatc 1800  
 ttgttggcgt gttgtcaatg ttgtgcttg aagaatcggc cttccggga ttttctgcgg 1860  
 tgtcacgggg ttttgcgttg accgtggagt tgggtcgggt tttgcgggca gctggtgca 1920  
 ctgcccgggt gtggttgcgt acgtgtggtg gcgtgtcggg gggagatgta ccggttcgtc 1980  
 cagcgcaggc cctggcggtg gggttggggc gtgttgtggg gttggagcat ccggactggt 2040  
 gggcgcgctt gatcgatatt ccggtcttgt tcgacgaaga cgctcaagag cggttgtcga 2100  
 ttgtgctggc aggtctcgat gaggacgagg tcgcatccg tctgacggc atgttcgcgc 2160  
 gtcggttggg acgccacact gtctcagctg atgtgaagaa ggcgtggcgc cccaggggat 2220  
 cgggtgctgg gacgggcggc acgggtggtt tgggggcgca cgttgctcgc tggctggccg 2280  
 acgccggagc cgaacatgtg gcgatggtga gtgcacgcgg cgagcaggca ccgagtgtc 2340  
 agaagttgcg gacggaactg gaggatctgg gtaccgggt gtcgatcgtg tcatgcgatg 2400  
 tgaccgatcg cgaggcgctc gccgaagtgc tgaaagccct tccggctgaa aaccggttga 2460  
 ccgcggtagt gcatgcggca ggcgtgatcg agactggtga tgcggcgga atgagcctgg 2520  
 ctgatttcga tcacgtgttg tccgcaaagg tggccgggtc cgcgaatctg gatgccttgt 2580  
 tggccgatgt ggaattggac gcgttcgtct tgttctcact ggtgtcagga gtttggggcg 2640  
 ctgggggaca cggggcttac gcagcggcga atgcctatct ggatgcgctc gcggaacagc 2700  
 gtcggtcgcg agggctggtc gcgactgcgg tggcctgggg gccgtgggac ggcgagggca 2760  
 tggcctccgg agaaacagga gaccagctgc gccgatacgg ctttcccca atggctccgc 2820  
 agcacgccat cgccggaatc cggcaggccg tggaacagga cgaaatttcc ctggtagtgg 2880

ccgatgtcga ttgggcaagt ttcagcgagg gattgctggc ggctaggccg cggccgctgc 2940  
tgaacgaact ggccgaggtc aaggaaactcc tcgtcgatgc ccagcccagag gcgggagtc 3000  
ttgccgacgc gtcgttggaa tggcggcagc gattgtccgc ggcaccgagg ccgacacagg 3060  
aacagctgat cctggagctg gtacgcggcg aaaccgctct ggtgctggga cccccgggg 3120  
cagcggccgt tgcacggaa cgagccttca aggacagcgg attcgactcg caggccgagg 3180  
tcgaactccg cgttcggctc aatcgagcta ccggcctcca gttgccatcg acaattatct 3240  
tcagccatcc cagcctgag gaactggctg cggagctgag ggcgaggctt cttcccagat 3300  
ccgcaggagc aggcattccc gaggaggacg aggcgcgaat cagagcggca ctgacgtcga 3360  
tcccgttccc ggccttgagc gaggcaggct tggtagagtc gctgctcgca cttgccggac 3420  
acccggctga ctccgggtat tcctcggagc atgcggccgc gacctcgatc gatgcgatgg 3480  
atgtagccgg cctcgtcgaa gcagcgtgg gcgaacgcga gtcctgagac cgcgcacctg 3540  
ggagatgacg gtgaccacca gttacgaaga agttgtcgag gcaactgcag catcgtcaa 3600  
ggagaacgaa cgcctccggc gcggcaggga tcggttctcc gcggagaagg acgatcccat 3660  
cgcgatcgtg gcgatgagtt gtcgttatcc cggtcaggtc tcctcgcggg aggacctgtg 3720  
gcaactggct gccggcgggtg tggacgcgat ctccgaagtt ccgggggagc gcggatggga 3780  
cctggatggc gtgttcgttc cggactccga tcgtcctggc acgtcgtatg cctgcgcggg 3840  
cgggtttctt cagggcgtgt cggagttcga cgcgggtttc ttcgggattt cgcgcgtga -3900  
ggcgctggcg atggatccgc agcagcgtt gctgctggaa gtcgcgtggg aggtcttcga 3960  
gcgggctggg ctggagcagc ggtcgacacg cggttccgcg gttggcgtgt tcgtcggcac 4020  
caatggccag gactacgcgt cgtggttgcg gacgcgcgcg cctgcggtgg caggatcatg 4080  
gctgacgggc ggtgcggcag cggttctttc gggccgggtt gcgtattcgt tcgggttcga 4140  
gggtcctgag gtgacgggtg atacggcgtg ttcgtcgtcg ttggtggcgt tgcacctggc 4200  
ggggcaagca ctgcgggccc gtgagtgcca ccttgccctt gccggtggcg tcacgggtgat 4260  
gtcgacgccg aaggtgttcc tggagttctc ccgccaacgg ggtctcgcgc cggatgggag 4320  
gtgcaagtcg ttcgcggcgg gtgcggatgg cactggatgg ggtgaggggt ccggactgtt 4380  
gttgcgtggg cggttgtcgg atgcccggcg gaatgggcat gaggtgctgg cggttgttcg 4440  
tggtagtgcg gtgaatcagg acggtgcgtc gaatggtttg accgcgcgca atggttcgtc 4500  
gcagcagcgg gtgattaccc aggcgttggc gagtgccggg ttgtcgggtg ccgatgtgga 4560  
tgctgtggag gcgcattggg cgggcaacgc gcttggtgat ccgacgagg cgcaggcgc 4620  
gatcgccacc tacggccgtg atcgtgatcc tggccggccg ttgtggttgg ggtcgtgcaa 4680  
gtcgaacatc ggtcatacgc aagcggcggc ggggtgtggc ggtgtgatca agatggtgat 4740  
ggcgatgcgg cacgggcagc tgccacgcac gttgcacgtg gaatcgccgt cgcgggaggt 4800  
ggattggtcg gcggggacgg ttcaactcct tacggagaac acgcccgtgc ccaggagtgg 4860  
tcgtgttcgt cgggtggggg tgcgtcgtt cgggatcagt ggtactaacg cgcacgtcat 4920  
cctcgaacag cccccgggag tgccgagtcg gtctgcgggg ccgggttcgg gttctgtcgt 4980  
ggatgttccg gtggtgccgt ggatggtgtc gggcaaaaca cccgaagcgc tatccgcgca 5040  
ggcaacggcg ttgatgacct atctggacga gcgacctgat gtctcctcgc tggatgttgg 5100  
gtactcgtg gcgttgacac ggtcggcgt ggatgagcga gcggtggtgc tggggtcgga 5160  
ccgtgaaacg ttgttggtgc gtgtgaaagc gctgtctgcc ggtcatgagg cttctgggtt 5220  
ggtgaccgga tctgtggggg ctgggggccc catcggtttt gtgttttccg gtcagggtgg 5280  
tcagtggctg gggatgggccc gggggcttta ccgggctttt ccggtgttcg ctgctgcctt 5340  
tgacgaagct tgtgccgagc tggatgcaca tctgggcccag gaaatcgggg ttcgggaggt 5400  
ggtgtccggt tcggatgcgc agttgctgga tcggacgttg tgggcgcagt cgggtttgtt 5460  
cgcgttgacg gtgggcttgc tgaagttgct ggattcgtgg ggggttcggc cgagtgtggt 5520  
gttggggcat tcggtgggag agttggcggc ggcgttcgcg gcgggtgtgg tgcgttgtc 5580  
gggtgcggct cggttgggtg cgggtcgtgc ccggttgatg caggcgttgc cgtctggcgg 5640  
tgggatgctg gcggtgcctg ctggtgagga gctgttgtgg tcgttgttgg ccgatcaggg 5700  
tgatcgtgtg gggatcgccg cggtaaacgc tgcggggtcg gtggtgctct ctggtgatcg 5760



ggatgtgctc gatgaccttg ccggtcggtt ggacgggcaa gggatccggt cgaggtggtt 5820  
 gcggtgtgct catgcgtttc attcgtatcg gatggatccg atgctggcgg agttcgccga 5880  
 attggcacga accgtggatt accggcgttg tgaagtgcg atcgtgtcga cttgaccgg 5940  
 agacctcgat gacgtggca ggatgagcgg gcccgactac tgggtgctc aggtgcgaga 6000  
 gccggtccgc ttcgcccagc gtgtccaggc gctggtcgag cacgatgtgg cactgttgt 6060  
 cgagctcggt ccggacgggg cgttgtcggc gctgatccag gaatgtgtcg ccgcacccga 6120  
 tcacgccggg cggctgagcg cgttcccggc gatgcgcagg aaccaggacg aggcgcagaa 6180  
 ggtgatgacg gccctggcac acgtccacgt acgtgggtgg gcggtggact ggcggtcgtt 6240  
 cttcgccggt acgggagcga aacaaatcga gctgccacc tacgccttc aacgacagcg 6300  
 gtactggctg gtgccatcgg attccggtga tgtgacaggt gccggtctgg ccggggcgga 6360  
 gcatccgctg ttgggtgctg tgggtccggt cgcgggtggg gacgaggtgt tgcgtaccgg 6420  
 caggatttcg gtgcggacgc atccgtgggt ggccgaacac cgggtgctgg gtgaagtgat 6480  
 cgttgccggc accgcgttg cggagatcgc cttgcacgcg ggggaacgtc ttggttgtga 6540  
 acgggtggaa gagtcaccc tggaaagcacc gctggtcctg ccggagcgcg gggcgatcca 6600  
 ggttcagctg cgagtgggcg cggccgagaa ttccggacgc aggcgatgg cgtgtattc 6660  
 acgccccgaa ggggcggcgg agcatgactg gacgcggcac gccacgggccc ggttggcgcc 6720  
 aggcgcggc gaggcggctg gagacctggc cgaactggcg gctcctggcg cgtgcccgtt 6780  
 cgacctcgac gaattctatc gggacctcgc agagcttggg ctggagtacg gccgatctt 6840  
 ccaagggctc aaggcggcct ggcggcaagg ggacgaggtg tacgcggaag ccgcgctgcc 6900  
 gggaaacggaa gattctggtt tcgggtgca tccggcactg ctggacgcgg ctctgcacgc 6960  
 aacggctgtc cgagacatgg atgacgcacg cttgccgttc cagtgggaag gtgtgtccct 7020  
 gcacgccaag gccgcgccgg ctttgcgggt ccgcgtggtc ccggctggtg acgatccaa 7080  
 gtccctgctg gtttgtgatg gcaccggctg accggtgatc tcggtgacc gactcgtatt 7140  
 gcggtcggct gcggcccggc ggaccgggtg gcgcccagag gcccatcaag ctcggttgta 7200  
 ccggttgagc tggccaacgg ttcaactgcc gacatccgct cagccaccgt cctgcgtgct 7260  
 tctcggcacc tcagaagtgt ccgctgacat acaggtgtat ccggacctcc ggtcgttgac 7320  
 ggctgcgttg gatgccgtg ccgaaccacc cggcgtcgtc atcgaccca cggcccccg 7380  
 cggtgagcga acagcggatg tccgggagac gactcggcat gactcgacc tggtaacagg 7440  
 ctggctttcc gatcagcgac tcaacgaatc ccgattgtc ctggtgacac agggagcagt 7500  
 ggcggtggag ccgggcgaac ccgtgaccga tctggcgag gccgcgctct ggggactgct 7560  
 gcggtcgacg cagaccgaac accctgatcg cttcgtctc gtcgatgtgc ctgagccgc 7620  
 gcaactctc cccgcgctgc cgggggtgct ggctgcggc gaacctcagc tcggttgctg 7680  
 acgtggcggc gctcatgcgc ccagactggc tggactggc agcgatgacg tcctgccgt 7740  
 gccggacggc accgggtggc gattggaggc cagcgcccg ggaagcctgg atgggttggc 7800  
 attggtggac gaaccgacgg ccacggcacc gctgggtgac ggtgaggtca ggattgcgat 7860  
 gcgcgcggcc ggggtgaact tccgggatgc gctcatcgcg ctcggtatgt atcccggtgt 7920  
 ggcacgctg ggcagtgagg gcgcgggggt cgtggtgagg accggccccg gcgtcaccgg 7980  
 cctggcacc gccgaccgcg tgatgggaat gatcccgagg gcgttcgggc cgtcgcgggt 8040  
 cgccgacat cgcattgtga cgaggattcc cgtgggttg agcttcgcgc gggccgcac 8100  
 ggtgccgac gtctttctca ccgcctacta cgcgctgggt gatctcgccg ggttgagacc 8160  
 aggggagtcg ttgctggttc attcggccgc cgggtgggtg gggatggccg cgatccaaact 8220  
 cgccaggcac ctcggtgcag aggtgtacg caccgctagc gaggacaagt ggcaagccgt 8280  
 ggagctgagc cgagaacacc tcgcttcgtc gcggacgtgc gatttcgagc agcagttcct 8340  
 cggggcaacc ggcggacgcg gcgtcgacgt cgtgctcaac tccctcgccg gggagttcgc 8400  
 cgatgcgtct ctgcgaatgc tgccgcgcgg tggccgttcc ctggagttgg ggaagacgga 8460  
 tgttcgtgac ccgctcgagg tcgccgatgc gcatccgggc gtgtcttacc aggttttcga 8520  
 taccgtagag gcaggccgc agcgaatcgg cgagatgctt cagagctgg tggagttgtt 8580  
 cgagggacgc gtgctggagc cctgcctgt cagcgcttgg gacgttcggc aggcgcccg 8640



ggcgctacgg	cacctgagcc	aagcgcgcca	tgtgggaaag	ctgggtgctca	ccatgcctcc	8700
ggtgtgggac	gccgcaggca	cggttctggt	taccggcgga	acgggagcac	ttggcgcaga	8760
ggtcgcccg	cacctcgtga	tcgagcgcg	ggtgcgaaac	ctggtcctcg	tcagcaggcg	8820
cggccccga	gccagtggcg	ctgctgagct	cgtggcgcaa	ctgacggcct	acggtgccga	8880
ggtttccttg	caggcttgcg	atgtcgccga	tcgtgagacc	ttggcgaagg	tgcttgccag	8940
catccccgac	gagcatccgt	tgaccgccgt	ggtgcacgcg	gctgggtgtc	tcgacgacgg	9000
agtgtccgaa	tcgctcaccg	tggagcggt	ggaccagggt	ctgcgcccga	aggtcgatgg	9060
cgcgcggaat	ctgctcgagc	tgatcgaccc	ggacgtggcc	ctcgtgttgt	tctcgtcggt	9120
gtcgggtgtg	ctcggcagcg	gtgggcagg	taactacgcg	gcggccaact	ccttcctcga	9180
cgcattggcg	cagcaaaggc	agtcgcgcg	cctaccgacg	agatcattgg	cctggggggc	9240
ctgggcggaa	catggcatgg	ccagcacctt	gcgcgaagcc	gagcaggatc	gattggcgcg	9300
atctgggttg	ctgccgatct	cgaccgagga	ggggttgtcc	cagttcgacg	ccgcgtgcgg	9360
cggcgcgcat	accgtggtgg	cgcgggttcg	attcagccgc	ttgtccgacg	ggaacgcgat	9420
caagttctcc	gtcctgcaag	gtttggtcgg	gccgcctcgc	gtcaacaaag	cggcgactgc	9480
ggatgatgcc	gagagcctcc	ggaaacggtt	gggacgcttg	ccggatgcag	aacaacatcg	9540
gattctgctg	gacctcgtcc	gcatgcatgt	ggcggcagtg	ctcggattcg	ccggttctca	9600
ggagatcacc	gcggacggca	cgttcaaggt	gctgggcttc	gactcgttga	ccgtgggtcga	9660
gttgcgcaac	cggatcaacg	gggcgacggg	gctgcgactg	cccgccaccc	tggtgttcaa	9720
ctacccgacg	ccggatgcgc	tcgccgcgca	cctcgtcacc	gcgctgtccg	cagaccgcct	9780
ggccgggaca	ttcgaggaac	tcgacaggtg	ggcggcgaac	ctgcccacgc	tgccagggga	9840
tgaggccacg	cgggcgcaga	tcaccacccg	gctacaggcg	atcttgacga	gcctggcgga	9900
cgtgtccggc	ggaaccggcg	gcggctccgt	gccggaccgg	ctcagatcgg	ccacggacga	9960
cgaacttttc	caactcctcg	acaacgatct	cgaacttccc	tgatgcctca	gccggagcct	10020
tcgcaacttc	ctggagggaa	acgccacatg	tcgaatgaag	agaagctccg	ggagtacttg	10080
cggcgtgcgc	tcgtggatct	gcaccaggcg	cgcgagcggc	tgacagaggc	ggagtcggga	10140
gagcgggaac	ccatcgcgat	cgtggcgatg	ggctgccggg	acccgggtgg	ggtgcaggac	10200
ccggaagggc	tgtggaaact	ggtgcgcctc	ggtggcgacg	ccatcggtga	attccccgct	10260
gatcgtggtt	ggcacctcga	cgagctctac	gatcccagac	cggatcagcc	cggaacctgc	10320
tacacccggc	acggcggtt	cctccacgac	gccggcgagt	tcgacgcggg	attcttcgac	10380
atcagcccc	gtgaggcgct	cgcgatggac	ccgcagcagc	ggctgctgct	ggaaatctcc	10440
tgggagaccg	tcgaatccgc	tgggatggac	ccgaggtcct	tgccggggag	ccgcaccggg	10500
gtgttcgcgg	gattgatgta	cgagggtctat	gacaccggcg	cccaccgggc	aggagaagggt	10560
gtcgaaggct	atctcggaac	cggcaatgcg	ggaagcgtcg	cctctggtcg	ggttgcgat	10620
gcgttcgggt	tcgagggcc	agcgggtgacg	gtagacacgg	cgtgctcgtc	gtcgttggtg	10680
gcgctgcatt	tggcgtgtca	gtcgttgccg	cagggcgagt	gtgatctggc	gctggccgggt	10740
ggagtgcagg	tgatgtcgac	gccggagagg	ttcgtggagt	tctcccgta	gcgtggtctc	10800
gcaccggatg	ggcgtgttaa	gtcgttcgcg	gcggctgcgg	atggaaccgg	ttggggtag	10860
ggtgccggtt	tggtgttgct	ggagcggctg	tcagacgcca	ggcggaacgg	gcacgggta	10920
ctggcggttg	ttcgtggtag	cgcgggtgaat	caggacggtg	cgtcgaacgg	attgacggcc	10980
ccgaacgggc	tggcccagga	gcgggtcatt	cagcagggtc	tcacgagtgc	ggggctgtcg	11040
gcgtccgatg	tggacgctgt	ggaggcgcat	ggaacgggta	cgcggcttgg	tgatccgatc	11100
gaggcgcagg	ctctgatagc	cgcctatgga	caggatcggg	accgggaccg	gccgctgtgg	11160
ttggggctcg	tcaagtccaa	catcggtcat	acgcaggcgg	ctgcggcgct	cgctggtgtg	11220
atcaagatgg	tcatggcgat	gcggcacggg	gagctgccgc	gcacgttgca	cgtggacgag	11280
ccgaattcgc	acgtggactg	gtcggctggt	gcggctccgac	tctgaccga	gaacatccgc	11340
tgccagggga	cgggtacgcg	ccgcgctgga	gtgtcgtcgt	tcggggtaag	cggtagcaac	11400
gcacacgtca	tcctcgaaca	cgaccgcctc	gccgtgaccg	agaacgagga	agcagcgacg	11460
tccccagcac	ctgggatcgt	gccctggggc	ttgtccgggc	ggtcgtcgac	ggcgtgcgg	11520

gcccaggccg aacggctgcg cgagctgtgc gagcagaccg atccccgacce cgtcgatgtc 11580  
ggttttctcac tggccgccac ggcacaggct tgggagcacc gagcgggtggt gcttggtcgg 11640  
gacagcgcta cgttgcgctc cgggcttggc gttgttgcca gcggtgaacc agcggctgat 11700  
gtcgttgagg ggagcgctct ggacggcgag gtcgtcttcg tcttccccgg tcagggtggt 11760  
cagtggggccg gtatggcagt cgacctgctg gacgcttcgc cgacgttcgc gcgccacatg 11820  
gacgagtgcg ccacccgcgct gcggaggtag gtggactggt cgttggtcga cgtgctgcgc 11880  
ggagcggaga actccccacc gctggaccgg gtggacgtgc tccagccccg gtccttcgcg 11940  
gtgatggtgt cgctcgccga ggtgtggcgt tcctacgggg tgaggccggc ggccgtcgtc 12000  
ggccacagtc aaggcgaaat cgccgcggcc tgcgcagccg ggggtgctgcc gctggaggat 12060  
gcggccaggc ttgtcgcatt gcgcagcaga gcgttggaagg gactttcggg gcgggggtggc 12120  
atggcgtcgc tggcctgccc tgcggatgag gtcgcggcat tgttcgcggg atcgggcccgc 12180  
cgtctggaag ttgcggcgat caacggcccc cgatcggtcg tgggtgcccg cgatctggaa 12240  
gcggtggacg aactgctggc agagtgcgt gaaaaggaca tgcgtgcacg ccgtatcccc 12300  
gtcgactacg cctcgcatte agcgacgtg gaggtggtt ggagccccgt gctggcggcc 12360  
gccgccgggg tgcgacaccg ggacggccag gtgcccgtgt ggtcgacggt gatcggcgac 12420  
tgggtggatc cggccaggct ggacggcgag tattggtatc ggaacctccg gcagccggtc 12480  
cggttcgaac acgccgtgca gggcctggtc gagcggggat tcggcctgtt catcgaaatg 12540  
gtgcgcatc cggtgctgac cagggcggtc gaggaaaccg gtgcgagtc ggagaccgcc 12600  
gtggcccgcg taggtacctt gcgacgtgac tcgggcggcc tccggagggt gttgcattcg 12660  
ctggccgagg cgtacgtgcg cggcgccacc gtggactggg ccgtggcgtt cgggggcgcg 12720  
ggccgacggc tggacctgcc gacctaccg ttccagcgcc agcggtagtg gctggacaag 12780  
ggagctgcct ccgacgaggc tcgtgcggtc tcggaccggc cggcgggctg gttctggcaa 12840  
gccgtggcgc gccaaagacct gaaaagcgtg tccgatgccc tcgatctcga cgcgcagcga 12900  
ccgctgagcg caacacttcc agccctgtcc gtctggcacc gtcaggaacg agaaagggtc 12960  
ttggcagacg gttggcggtta ccgagtcgac tgggtacggg tggccccga gccggctcgg 13020  
agaacgcggg aaacctggct cctggctcgt cccccggcg gcacgagga agcgtggtc 13080  
gaacggctga cggatgcgtt gaacacgcga gggatcagca ccctgcgcct cgacgtgcca 13140  
ccggcggcga ccagtggcga actcgcaacc gaactccgcg ccgcagccga cggtagcccg 13200  
gtgaaggcaa tcctgtcgtt caccgcgtt gacgagcgac cccaccccga atgcaaggac 13260  
gtcccagcgc ggattgcctt gctgctgaac ctggtcaagg cgtcgggtga agccgacctc 13320  
agaattcctc tgtggaccat cagcgtggt gcggtcaagg caggccccgc agatcggtg 13380  
ctgcgcccga tgcaggcgca agcatgggt ctggggcgag tagccgcact cgaacacccc 13440  
tagcgtggtg gtgggtgat cgacctgccg gattcgctgg acggcgacgt cctcacgagg 13500  
ctggggcgaag cgtcaccaa cggcttggcg gaagaccaac tggcgattcg ccagtgggc 13560  
gtgctggccc ggcgactggt acccgccccg gcgaatcagc ccgctggacg taagtggcgc 13620  
ccccgagga gcgcgtgat cagggcgga ctcggcgcgg tgggcgcaca ggtggcgagg 13680  
tggttggccg aaatcggagc cgagcgaatc gtgctacca gtcgacggg caaccaagca 13740  
gcaggcgccg ccgagctgga agccgaactc cgggccctt gagcgcaagt gtccatcgtg 13800  
gcttgcgacg tgaccgatcg tgcgagatg tccgcaactac tggccgagtt cgacgtcacc 13860  
gcggtgttcc acgcggccgg agtcggtcgg ctgctgccgt tggcgagac cgaccagaac 13920  
ggcctggccg aaatatgcgc ggcgaaggte cgcggcgctc aggtgctgga cgaactgtgc 13980  
gacagcaccg atctcgatgc ctctgctctg ttctctcctg gtgcccgggt atggggcggg 14040  
ggcggctcagg gcgcttacgg cgcggcgaa gcattcttgg acacactcgc cgaacaacgc 14100  
cgagcacgcg gtctgccggc aacctcgatc tcctggggca gttgggcccg cggcggcgatg 14160  
gccgacggcg cggcgggcga acacctgcgg cgacgcggga tacgtccgat gccggcgcg 14220  
tcggccatcc tggctctgca ggaagtactt gaccaggatg agacgtgcgt gtcgatcgct 14280  
gatgtggact gggaccgatt cgttcccacg ttcgccgcga ctcgcgccac ccggttgttc 14340  
gacgaagtgc cggcggcgag aaaggcgatg cccgcgaatg ggccggcaga accaggcggc 14400

tcgccgttcg cccgcaatct cgcggagctg ccggaagccc aacgacgcca cgaactggtg 14460  
 gatctggtgt gcgcccaggt ggcaaccgtg ctccggcacg gcagtcgcga ggaagtccag 14520  
 cccgagcggg cgttccgcgc gctcgggttc gactccctca tggcgggtga tctgcgcaat 14580  
 cgtttgacca ccgccaccgg gttgcgcctg ccgaccacaa ccgtcttcga ctaccggaat 14640  
 ccggccgcct tggccgctca cctgctcgag gagctggtgg gtgatgtcgc gtcggctgcg 14700  
 gtgaccgctg ccagcgcgcc cgcgagtgac gaaccgatcg cgatcgtcgc gatgagctgc 14760  
 cggtttccgg gtggcgcgca ctcccggaac gacctgtggc ggctggtcgc cgccggcacg 14820  
 gaggtgatcg gcgagttccc ctccgaccgg ggctgggatg cggaaggcct ttacgatccg 14880  
 gatgcttcca ggcctggaac gacgtatgcg cggatggcgg gattcctcta cgacgccggt 14940  
 gagttcgatg ccgacctgtt cggcatcagc ccacgtgagg cgttggcgat ggatccgcag 15000  
 cagcggttgg tgctcgaaat cgctgggaa gccctcgaac gggccggaat cgatccgttg 15060  
 tccttgaagg gcagtggggt cggcacgtac atcggcgctg gaagccgtgg gtacgcgacg 15120  
 gatgtgcggc agtttcccgga ggaggcggag ggctacctgc tgacgggtac ctccggccagt 15180  
 gtgctgtcgg gtcgggtcgc gtattcgttt ggtttcgagg gtcctgcggg gacggtggat 15240  
 acggcttgtt cgtcgtcgtt ggtggcgttg catctggcgt gccagtcgtt gcgttcgggc 15300  
 gagtgtgatc tggcggttggc cgggtggtgt accgtgatgt cgacgccgga gatgttcgtg 15360  
 gagttctccc gtcagcgcgg tttggcgccg gatgggcggg gcaagtcgtt cgcgagagac 15420  
 gcggacggca ccggctgggg cgaaggcgcg ggccgttgtt tgctggagcg gttgtcggac 15480  
 gccaccgga atgggcatcg ggtgttgggc gtggttcgtg ggtcagcggg gaatcaggac 15540  
 ggcgcctcga acggactggc ggcgcggaac ggtccgtcgc agcagcgggt gatcaaccag 15600  
 gcaactcgcga atgcggctct ttcggcgtcc gatgtggatg cgggtggaggc acatggcacc 15660  
 gggaccaggc tgggtgatcc gatcgaggcg caggcattga tcgcaacgta tgggcaggcc 15720  
 cgggagcggg atcggccctt gtggctgggg tcggtcaagt cgaacatcgg tcatacgag 15780  
 gccgcggcgg gtgttgccgg tgtgatcaag atggtgatgg ccatgcggca cgggcagctg 15840  
 cccgcctcgc tgacgcgga tgagccacg tcggaggctg attggtcgtc gggggcggtc 15900  
 cggctcctcg ccgaacaggt accttgccg gagtctgacc gtgttcgtcg ggtgggggtt 15960  
 tcgtcggttcg ggatcagcgg caccaacgca catgtgatcc tcgaacaagc tacgaatgcg 16020  
 ccagatagta cagcggagac ggacaaaaca gaatccggat ctactgtcga tattccggtc 16080  
 gttccctggt tgggtgtcgg aaagacgac gattccctgc ggggacaagc cgaacgagtc 16140  
 ttgtctcagg tcgagtcctg gccggagcag cgttcgctgg atgttgctta ctcgcttgct 16200  
 tctggccgag ccgcgctgga tgaacgcgct gtcgtgctgg gtgcggaccg cggtagctg 16260  
 gttgctggac tggcggcgtt ggccgcgggt caggaggctt ctggggtgat cagcggaaact 16320  
 cgtgcttctg ctcggttcgg gttcgtgttc tcggggcagg gtggtcagtg gttggggatg 16380  
 ggcagagcgc tctactcgaa gtttccggtg ttcgctgctg cgtttgatga ggcttcgcc 16440  
 gagttggagg cacatctggg ggaagaccgc cgggttcggg atgtggctct cggttccgat 16500  
 gcgcagctgc tggatcagac gctgtgggcg cagtcgggtc tgttcgcgct gcaagccggc 16560  
 ctcttggggc tgctgggttc gtggggcggt cggccggatg tggtagtggg gcattcggtc 16620  
 ggggagttgg ccgccgcgtt tgccgctggc gtgttgctgt tgccggatgc ggctcgggtt 16680  
 gtggccgcgc gcgcccggtt gatgcaagcc ctgccctctg acggcgcgat gttggcggtg 16740  
 gctgctggtg aagaccttgt tcggccattg ctggccggtc gggaggagtc cgtgagcgtc 16800  
 gccgcgctca atgcccccg ttcgggtggt ttgtcgggag atcgggaggt gctggccagc 16860  
 atcgtcggcc ggctgaccga gctccgagtc cggacgcggc gcttcggggg ctcccatgct 16920  
 tttcattcgc accggatgga cccgatgttg ggcgagttcg ccagatcgc cgagtctgcg 16980  
 gagttcggta agccaacgac accgcttggt tcgacgttga cgggtgagct cgacagagcc 17040  
 gcggaaatga gcacaccagg gtattgggtg cggcaggcgc gtgaaccgtt ccgtttcgcc 17100  
 gacggtgtcc aggccttggc agcgcagggc ataggcacgg tcgtcagact cggcccgagc 17160  
 ggaacgctgg cggcactggt tcgggagtg ggcaccgagt ccgatcgggt tgggcggatt 17220  
 tcgtcgatcc cactgatgcg cagggagcgg gacgagacc gttcgggtgat gacagccctg 17280

gcgcacatccc acacccgtyg tgggtgaggtg gactggcagg cgtttttcgc cgggtaccggc 17340  
 gctaggcagc tcgagttgcc aacgtatgcc ttccaacgac agcactactg gatcgagtcc 17400  
 agtgcgcggc cagcacgcga ccgcgcagac atcggcgagg tggcggaaca gttctggacc 17460  
 gcggttgacc aaggcgatct ggcaacgttg gtcgcgcgtc tggatcttgg ggcggacgac 17520  
 gacacatgcg catcggttag cgatgtattg ccggcggttg cctcctggcg aagcggactc 17580  
 cgcaaccggt cgctcgtcga ttctgcccgg taccgaatca gttggcattc ctctcgggag 17640  
 gtgcccggccc cgaagatttc cggtagcttg ctggttggtc tgcggggtgc tgcggatgac 17700  
 ggattggtca cggctttgac gagttcactg gtcggaggcg gcgcccaggc cgtccggatc 17760  
 ggctgtccg aagaggaccc gcaccgcgag gacgtcgac agcggctggc caatgcgctg 17820  
 acggatgccg gtcaactcgg tggcgtgctt tcgctgttgg ggctcgatga atcgctgct 17880  
 ccgggattct cctgcttgcc aactggtttc gcgctgactg tgcagcttct gcgggccttg 17940  
 cgaaggccg acgtcgaggc gcctttttgg gcggtgacgc gcggcgccgt cgcgttgaa 18000  
 gatgtacgcg tgtctccgga gcaggccctg gtctgggggc tgcgtgctgt cgcgggactg 18060  
 gagcaccggg agttctgggg tggcttgatc gacctgccat cggactggga cgaccgattg 18120  
 ggtgcccggg tggcggtgt gtggcggtg ggtggcgagg atcaagtgc cattcgccgt 18180  
 ggtggtgtgt tcgtgcggcg gttggaacgc gctggtgctg cgggtgccgg gtcggtgtgg 18240  
 ggtcctcggg ggacggtgtt ggtgacgggt ggtacggcg gtttgggggc gcatgttgcc 18300  
 ggtggttg ccggtgccgg ggctgagcac gtggtgttga ccagccgtcg aggagcggac 18360  
 gctccggggc ctggggaatt gcgggcggag ctggaggcgc tgggtgctcg ggtgtcgatt 18420  
 gtgcccgtcg acgtggctga tcgtgacgca gtggctggag tgttggcagg gatcggtggg 18480  
 gagtgtccgc tgcactgcgt ggtacacgcc gccggggctg gcgaggcggg cgacgtagt 18540  
 gagatgggtt tggcggtatt tgcagcggtg ttgtcggcga aggtgcgtgg tgcggcgaat 18600  
 ctggacgagt tgcgtggcga ctcgagctg gatgcgtttg tgatgttctc ctcggtgtcg 18660  
 ggggtgtggg gagccggcgg acagggtgcg tatgcggctg cgaacgccta cttggatgcg 18720  
 ttggccgagc agcgtcgggc gaggggattg gtcgggaccg cggttgcgtg gggaccgtgg 18780  
 gccggtgacg gcatggccgc cggcgaaacc ggccgacagc tgcaccggtt gggcctggcg 18840  
 tcgatggaac cgagcgcggc gctgctggca cttcagggtg cattggaccg cgatgagacc 18900  
 tccctcgtcg tggccgatgt cgattgggca cggttcgccc cagccttcac ctcggcacgt 18960  
 cgacgcccgc tgcgtggacac catcgacgag gcccgagccg cattggaaac caccggcgaa 19020  
 caagcgggca caggcaaacc cgttgagctg acgcaacgcc tggccggact gtcgcggaag 19080  
 gaacgcgacg atgcggtatt ggatctggtg cgggcggaga cggcggtgt gctgggacgc 19140  
 gacgatgcca cggccctggc gccatcgcg ccgttccagg aactcggatt cgactccttg 19200  
 ctggcggtgg agctgcgcaa ccggctgaac accgccaccg ggatccagct gcccgccagc 19260  
 cgatgttctg actaccccaa tgccgagtcg ctgtcgcgtc acctctgcgc cgagcttttc 19320  
 ccaacggaga ctaccgtgga ctcgccctt gccgagctcg atcgaatcga gcagcagctc 19380  
 tcgatgtca ccggcgaagc gcgggcacgg gaccgaatcg cgacacgact gcgagccctc 19440  
 cagcagaagt ggaacagcgc agctgaagta ccgaccggag ccgatgtcct gacacgctc 19500  
 gattcggcga cgcacgacga gatattcgag ttcatcgaca acgagctcga cctgtcctga 19560  
 gcagttcctg cggaaacttca agcgcgaaa tcgggtggaa atcacaatgg ccaatgaaga 19620  
 aaagctcttc ggctatctga agaaggtaac tgcggacctg catcagacc gccagcgct 19680  
 gctcgcggcc gagagccgga gtcaggagcc gatcgcgatc gtctcggcga gctgcggact 19740  
 gcccgcggc gtcgactctc ccgaagcgt ctggcaactc gtgcgactg gcaccgacgc 19800  
 catctcgag ttccccgcg accggggctg ggatctcggc cggttgtacg atcccgacc 19860  
 gaaccaccag ggaacgtcgt acacgcgggc cggcggtttc ctgcaggag cgggcgattt 19920  
 cgaccccgcc atgttcggga ttctgcgcg tgaggcgtt gcgatggacc cgcagcaacg 19980  
 gttgttgcg gagctgtcct gggaggccct cgaacggcg ggcatagacc cgacatccct 20040  
 gcgcggcagc aagaccggtg tcttcggtg tgcacgcc caggagtac ggccgtcctt 20100  
 gcaggagatg agccgaaacg ctgggggttt tggactcacc gggcggtatg tgagtgtggc 20160

gtcgggtcgg gttgcgtatt cgtttggttt tgagggctct gcggtgacgg tggatacggc 20220  
 gtgttcgtcg tcgttggtgg ccctgcattt ggcggtgcag tcgttgcggt ccggcgcaatg 20280  
 cgatctcgcg ctggccggcg gtgtgacggt gatggcgaca ccggcgacgt tcgtggagtt 20340  
 ctcccgtcag cgtggttttg ctccggacgg gcggtgcaag tcgttcgcgg ctgccgcgga 20400  
 tggcaccggg tggggtgagg gtgccggtct ggtgttgctg gagcggttgt cggatgcgcg 20460  
 gcggaatggg cagcaggttc tggcggtggt gcggggtagc gcggtgaacc aggacggcgc 20520  
 gtcgaatggt ttgactgcgc cgaatggtcc gtgcgcagcag cgggtgatca cccaggcggt 20580  
 ggcgagtgcg gggctgtcgg tttccgatgt ggatgcggtc gaggcacatg ggaccgggac 20640  
 cacgttggtt gatccgatcg aggcacaggc cctgatcgcc acgtacgggc agggccggga 20700  
 gaaggatcgg ccgttggtgt tggggtcggc caagtccaac atcggtcaca cgcaggcggc 20760  
 cgctggcggt gccggcgta tcaagatggt cttggcgatg cggcacgggc agctgccccg 20820  
 cacgttgcat gtggatgagc ccacgtcggc ggtggactgg tcggcggtt cggctccggct 20880  
 tctcacggag aacacgccct ggccggacag tggctcgtct tggcggttg ggtgtcgtc 20940  
 gttcgggatc agcggcacca acgcacatgt gattctcgaa cagtctccag tcgagcaggg 21000  
 cgaaccggcc gggccggtcg aaggcgagcg ggaaccggat gttagcgtcc ccgtggtgcc 21060  
 ttgggtgctg tcgggtaaga caccggaggc tgccggggcg caggccgaac ggggtcattc 21120  
 gcatatcgag gaccggccgg ggtgtcgcgc ggtggatgtg gcgtattcgc taggaatgac 21180  
 acgcgcggcg ctggatgaac gcgcagtggt gttgggctcg gaccgtgccg cgctcctgac 21240  
 cgggttgagg gcattcgcgc acggctgcga tgcgccgaa gtggtttcgg ggtctgtggg 21300  
 gcttggtggc cgcgtcgggt tcgtgttctc gggtcagggt ggtcagtggc cggggatggg 21360  
 ccgggggctc tactcgggtt tcccggtgtt cgccgacgcg ttcgacgagg cttgcgcgga 21420  
 gttggatgca cacctgggac aggaactgcg ggttcgggat gtggtgttcg gttcgcaagc 21480  
 gtggttgctg gatcggacgg tgtgggcgca gtcgggtttg ttcgcgttgc agattggctt 21540  
 gctgcggctg ctgggttcgt ggggtgttcg gccggatgtg gtgttggggc actcgggtggg 21600  
 tgagctggct gcggtgcag cggctggtgt gttgtcgtt tcggaggccg cgcggttgg 21660  
 ggcgggtcgc gcccggttga tgcaggcggt gccttctggt ggtgccatgc tcgcggctgc 21720  
 tacgggtgag tttcaggctc atcctctgct ggatggggtg cgggaccgga tcggtatcgc 21780  
 ggcggtgaat ggcccgaat cggttgtgct ctctggtgac cgcgagctgc tcaccgagat 21840  
 cgctgatcgg ttgcacgac aggggtgccc gaccgggtg ttgcgggtgt cgcgtgttt 21900  
 ccattcgcgc catatggagc cgatgctgga ggagtgcgc cagatctccc gaggccgcga 21960  
 atatcacgca ccggaactgc cgatcatctc gacctgatc ggtgagctgg acggtggtcg 22020  
 agtgatgggc actcccagat actgggtgcg tcagggtcgt gagcccgctc gtttcgccga 22080  
 ggggtgccag gcgcttgctg gtcagggtgt cggcacgatt gtcgaattgg gtccggacgg 22140  
 ggcggttgct acggttggtc aggagtggtt ggcggaatcc gggcggttg cgggatccc 22200  
 gctgatgcgc aaggaccgcg acgaggcgcg aaccgtgctg gcagcttttg cgcagatcca 22260  
 caccgtggt ggtgaggtgg actggcggtc gtttttcgcc ggtaccgggg cgaagcaagt 22320  
 cgacctgcc acctacgcct tccagcgga gcggtactgg ctggcatcca ccgggcgtgc 22380  
 gggtgacgtg accgccgccg gattggccga ggcggaccat ccgctgctcg gtgcggtggt 22440  
 tgcgttgga gacggcgaag gtgtggtgct gaccggctcg ttgacagcgg gttcgcattc 22500  
 gtggttgctc gatcaccggg tgcggggcga aatcgctgct cccggcaccg cgatcgtcga 22560  
 gctggtgtgg cactcgggcg agcgccctcg ttgtggccgg gtggaagaac tggcttttga 22620  
 agcgcacctg atcctgccgg atcatggagc ggtccagggt cagggtgctg tgggaccgcc 22680  
 cggggaatcc ggagcccggt cgggtggcgt ctactcctgt cctggcgagg cgatcgaacc 22740  
 cgagtggaag aagcacgcga cggcggtgct tctcccaccc gtggccgccg agaaccatga 22800  
 gctgaccgca tggcccccg agaatgcgac cgaaatcgat gcagacgggg tctacgcatt 22860  
 ccttgaaggg caggttttcg cgtacggacc ggcctttaga tgtctgcgcg gtgctggtgg 22920  
 acgaggcggg gaggtgttcg ccgaagtgc attgccggat gacatgcagg cgggggtcga 22980  
 tcgattcggc gtccaccccg cggtgctgga cgcggttctg catgccgccg cagccgagac 23040

gtcggtggtc cagagcgagc cgcgggtgcc gttctcgtgg cgtggggcgg aacttcgcgc 23100  
 cactgaaagc gcggtggtgc gggcgcgccct ctcggtgact tcggatgacg aactgtcgtt 23160  
 ggtcgcagtg gacccggctg gccgattcgt ggccacggtt gattcgctgg tgacccgacc 23220  
 gatctcccgg cagcaggtga ggtctggcgc gatcgggtgat tgcctgttcg aggtggagtg 23280  
 gcaccggaag gcgttggttg gaacaacgcg cggcgacgac cttgccatcg tcggtgacgg 23340  
 tcccagtttg ccggaatcgg tgcgcgcaac cgcacggttc gcgaccctgg atgagttccg 23400  
 tgcggccgtg gactcggacg ttctgcccc gggttcgggtg ttggtcgcag ctatgtcggc 23460  
 cgaagaggtc gaggggtgat ccctgccgtc gcgcgccccaa gagtcgacct ccgatctgct 23520  
 ggctctcgtg cagtcgtggc ttgaggacga gcggttcgcc gaatcccagc tcgtggtcgt 23580  
 cacgcgtgca gcggtgtcgg ccgactcggg ttccgacgtc gcggacctgg tgggtgcgtc 23640  
 gtcgtggggg ttgttgagtt cagccagtc ggagaaccgc ggtcgttcg tgctggtgga 23700  
 cgtggacggc acacctgagt ggtggcaggc gttgccggcc gccgtgcgag caggagaacc 23760  
 gcagctggca cttcggcgcg gcgtggcgct ggtgcctcgg ttggcgcgac tcacggtgcg 23820  
 cgaggagggc tctcccccgc aactcgacac ggacgggacc gtcctcatca cgggtggcac 23880  
 cgggtcgttg gggggagtgg ttgcccgtca cctggtggag gagcacggga ttccggtttt 23940  
 ggtgttgga ggcgcgcgtg gctggaatgc gcctggagtc cacgagttgg tggatgagct 24000  
 ggcgcgcgcg ggcgcgcgtg ttgaggtggt ggcttgcat gtggctgacc gcaccgatct 24060  
 ggagcacgtg ctggccgcca ttccggtcga ctggccgctg cgggggatcg tgcataccgc 24120  
 tggggtgctg gccgacggag tgatcgggtc cttgtcggcg gcggatgtgg gcacggtgtt 24180  
 tgccccgaag gtgacggggg catggcatct gcacgagttg acccgcgatc tggatctgtc 24240  
 gttcttcgtt cttttctctt cttctccgg gattgcgggt gccgcagggc aggccaacta 24300  
 cgcggcgcg aacacgttcc tggatgcatt ggcgcgttat cgcggggcgc gtgggtgcc 24360  
 tgggttgctg ttggcgtggg gactgtgggc gcaaccacgc ggtatgacga gtggcttga 24420  
 cgcggcgctg gtggagcggg tggcgcgac gggcatcgca gaactttcca cggaggatgg 24480  
 actcgcctg ttgatgccg cgttcgcgaa ggaccgggtt tgcgtcgttg ccgctcgatt 24540  
 ggacagggcg ctgctggtcg ggaacggacg atcgacgcg attccggcgc tgttgagcgc 24600  
 gttggttctt gttcgcggcg gtgtggcgag gaaaacagcc aattctcagg ccgcggatga 24660  
 ggacgcactg ttgggttttg tgcgggagca cgtttcggcc gtgctgggtt attcgggtgc 24720  
 ggtcgaggtt gggggcgacc gtgctttccg tgatctgggt tttgattcgt tgtctggcgt 24780  
 ggagttgcgg aaccgccttg ccgggtgct ggggtgctg ttgccggcga ctgcggtgtt 24840  
 cgactatccg acgcgcggg cgtggtgcg tttcctgcat caggaaactg caggcgaggt 24900  
 cgcgtccacg tcgacgccg tgaccagggc agcgagtgc gaagaggatc ttgttgcat 24960  
 gtcgggatg ggatgtcgtt ttccgggtg ggtgtcgtcg ccggaggagc tttggcggct 25020  
 ggtggccggc ggcgtggatg cgggtggtg gttcccagac gatcgcggct gggatctcgc 25080  
 ggcgtgtac gatcctgate ccgatcgtct cgggacctcg tatgtgtgtg agggcgggtt 25140  
 tctgcgggac gcggcgaggt tcgatgctga catgttcggc atcagcccgc gtgaggcgtt 25200  
 ggcgatggat ccgcagcagc ggttgctgct ggaggtcgcc tgggaaacct tggagcgggc 25260  
 tgggatcgat ccgttctcgt tgcacggcag ccggaccggt gtgttcgcgg gcttgatgta 25320  
 ccacgactat ggggcccgat tcattaccag agcaccggag ggcttcgaag ggcacctcgg 25380  
 gacgggcaat gcggggagcg tgctgtcggg tcgggttgcg tattcgtttg gtttcgaggg 25440  
 tctgcgggtg acggtggata cggcgtgttc gtcgtcgttg gtggcgttac acctggcggg 25500  
 tcaagcactg cgggcccgtg agtgcaatt cgccttgcc ggtggcgta cgggtgatgtc 25560  
 gacgccgacg acgttcgtg agttctcccg tcaacgggtt cttggctcgg atgggcgggtg 25620  
 caagtcgttc gcggcgcccg cggatggcac cgggtggggc gagggtgccc gtctggtgtt 25680  
 gctggagcgg ttgtcggatg cccggcgcaa tgggcaagag gttctggcgg tgggtcgggg 25740  
 tagcgcgggtg aaccaggacg gcgcgtcgaa tggttgact gcgccaatg gtccgtcaca 25800  
 gcaaagggtg atcaccagc cactcacgag tgccgggctg tccgtgtccg acgtggatgc 25860  
 tgtggaggcg catgggacgg gcacgcggct tgggtgatccg atcgaggcgc aggcgttgat 25920

cgctacgtac ggccgggata gtgatcccg tggccggtg tggctggggt cggatgaagtc 25980  
 gaatatgtgt cacaccagc cggcggcggg tgtcgctgtt gtgatcaaga tggatgatggc 26040  
 gatgcggcag ggggagctgc cgcgcacgtt gcacgtggac gagccctccg cgcaggtgga 26100  
 ctggtctgcg ggcacggtcc aactcctcac ggagaacacg ccctggcccg acagcggctcg 26160  
 tcttcgcccgg gcgggctgt catcggttcg gatcagtggt accaacgcgc acctgatcct 26220  
 tgaacaacct ccgcgagagt cgcagcgtc aacagagccg gattcgggtt ctgtccgcga 26280  
 ttttcgggtg gtgccgtgga tgggtgctggg caaaacaccc gaagcgctat ccgccagggc 26340  
 agatgcattg atgtcctact tgagcaatcg cgttgatgct tccccgcgag atatcggtta 26400  
 ttcgcttgcg gtgaccgctc cggcggttga ccaccgcgt gtcgtgctgg gtgcggatcg 26460  
 tgccgcgttg ctgccgggtc tgaaagcgct ggcggttagt aatgacgctg ccgaggtgat 26520  
 caccggcact cgtgccgctg ggcgggtcgg attcgtgttc tccggtcaag gtggtcagtg 26580  
 gcccgggatg ggaagcgggc tccactcggc gtttcgggtg ttcgccgacg cgtttgacga 26640  
 agcctgctgc gagctggatg cgcactcctg gcagatggcc cggctacgag atgtgttgct 26700  
 cggttcggat acgcaacttc tggaccagac cttgtgggcg cagccgggccc tgttcgctgt 26760  
 gcaagtgcga ctctgggagt tgttgggttc gtggggtgct cggcccgctg tgggtgctggg 26820  
 ccaactcggtc ggtgagctgg cggcggtcgt cgcggctgga gtgttgctgt tgcgggatgc 26880  
 ggtcggctg gtggcgggccc gtgccgggtt gatgcaagcc ctgccaaactg gcggtgccat 26940  
 ctcgctgctg gctgctggag aggagcagct gcgccggtt ctggccgact gcggtgatcg 27000  
 tgtggggatc gccgcggtca acgctcccg gtccggtggt ctctccggtg atcgggatgt 27060  
 gctcgatgac attgccggtc ggctggacgg gcaagggatc cggctccaggt ggttgccgggt 27120  
 ttcgatgctg tttcattcgc atcggtatga tccgatgctg gcggagttca ccgaaatcgc 27180  
 ccggagcgtg gactaccggt cgtcagggct gccgatcgtg tcgacgttga cgggtgagct 27240  
 cgatgaggtc ggcatgccgg ctacgccgga gtattgggtg cggcaggtgc gagaacccgt 27300  
 ccgcttcgcc gacggtgttg ctgcgctcgc ggctcacggt gtgagcaccg tcgtcgaggt 27360  
 cggctccgat ggggtgttgt cggcgctggt gcaggagtgc gcggccggat ccgatcaggg 27420  
 cggacgggtg gccgcggttc cgtcatgctg cagcaatcgc gacgagggcg acacggtgac 27480  
 aacggcattg gcgcagatcc atgtgctgtg tctgaggtg gactggcggt cgtttttcgc 27540  
 cggatccggg gcaaagcagg tcgagctgcc cacgtatgcc ttccaacgac agcgggtactg 27600  
 gcttgactca ccattccgaac cggctcgggca atccgccgat cccgcgcgcc agtcgggctt 27660  
 ctgggaactc gtcgagcagg aagatgtcag cgcgctcagc gccgctctgc acattaccgg 27720  
 cgatcacgac gtgcaggcgt ccctggaatc ggtggttccg gtccctcctt cctggcatcg 27780  
 ccggtaccgc aacgaatccc tgggtcacca gtggcggtac cggatttctt ggcatgagcg 27840  
 gcagatttg ccagaccctt cgttgctggg gacatggctc gtcgtcgtgc cggaggggtg 27900  
 tccggcgagt cggcaagttc tgcgtttcaa cgagatgttc gaggaacggg gttgcccggc 27960  
 agttctgttc gagctcgcgg ggcacgacga ggaagccctg gcgcaacgat tccgctcgtt 28020  
 gcctgttgcg tcagggggaa taagcggcgt gttgtccttg ctggcgctgg atgaatcgcc 28080  
 gtccctcgcc aacgctgctt tgccgaatgg cgcgctgaac tcgttggtac tgcgtcgagc 28140  
 tctgcgggccc gcggatgtgt cggcgccatt gtggttggtg acgtgtggtg gtgtcgcggt 28200  
 cggggatgtg ccggtgaacc cggggcaggc gctggtgtgg ggactgggtc gcgtcgtcgg 28260  
 tctggagcat ccggcctggt ggggtggcct ggtcgacgtg ccgtgcttgc tcgatgagga 28320  
 cgctcgagaa cgcttgctgg tcgtgttggt aggtcctggc gaggacgaga tcgcggtacg 28380  
 tcccgggtgt gtgttcgtgc ggcggttggg acgcgctggt gcggcgctcg gtgcccgggtc 28440  
 ggtgtggcgt cctcggggga cgggtgttgt gacgggtggt acgggcggtt tgggggcgca 28500  
 tgttgcccgg tgggtggcgg gtgccggggc tgagcatgtg gtgttgacca gccgtcgagg 28560  
 cgcggcggtc ccgggcgctg gagatttgcg ggcggagctg gaggcgctgg gcgctcgggt 28620  
 ttcgatcacg gcctgcgacg tggccgatcg tgacgctttg gccgaagtgt tggcgacat 28680  
 tccggatgat tgcccgtgta ccgcggtgat gcatgcggcg ggggtcgttg aagtcggcga 28740  
 cgtggcgctg atgtgtttga ccgacttcgt tggggtgctg tcggcgaggg caggtggtgc 28800



ggcgaatctc gatgagttgc tcgccgatgt cgagctggat gccttcgtgc tgttctcacc 28860  
 cgtctcgggt gtgtggggtg ctggcgggca gggcgcttat gcggcggcga atgcctactt 28920  
 ggatgcgttg gcgcagcagc gtcgggcaag ggggttgggtg gggactgcgg ttgcgtgggg 28980  
 cccgtgggccc ggtgacggaa tggccgcagg tgaaggcggg gcacagctgc gccggggccg 29040  
 cctggtgcca atggctgcgg atcgggcgtt gctggcactt cagggcgcat tggatcgtga 29100  
 cgagacatcc ctggctcgtg ccgatatggc gtgggagagg ttccgcccgg tttcgcccat 29160  
 gtcccgtcgg cgtccgctgc tcgacgagct gccgaagca cagcaggcgt tggcggatgc 29220  
 ggagaacacc actgatgctg cggactcggc cgtcccgtca ccgcggctcg cgggcatggc 29280  
 agccgccgaa cgccgcgcgg cgatgctgga cctggtgctg gcggaggcct cgattgtgtt 29340  
 gggacacaac gggctctgacc cagttgggtc cgaccgggcg ttccaggagc tcggatttga 29400  
 ttcgctgatg gccgtcgaac tgcgcaacag gttgggagag gcaacaggat tgagtctgcc 29460  
 ggccacgttg atcttcgatt atccgagccc atccgcgctg gctgagcagc tggtcggcga 29520  
 gctggtggga gcgcagccc cgaccacggt cgtggccggg gccgatccag tggatgatcc 29580  
 ggttgtcgtg gtcgcgatgg gatgccggtg tccgggcgac gtctgctcgc ccgaggagct 29640  
 gtggcagctg gtttctcggg gacgtgatgc ggtatcgacg ttcccctcg atcgggggtt 29700  
 ggactgcaac acgttggtcg acccggtacc ggatcgggca ggcagtacct atgtgcgaga 29760  
 aggtgccttc ctgaccggtg ctgatcgggt cgacgcgcgg ttcttcggca tcagccctcg 29820  
 cgaggcgcg ccaatggatc cgcagcagag gttgttgctc gaagtggcgt gggaggtttt 29880  
 cgaacgagca ggaatcgctc cgctgtcgtt gcggggtagc aggaccggtg tgttcgcggg 29940  
 gaccaatggg caggaccacg gtgcgaaagt ggctgcgcgc ccggaggcgg cgggtcacct 30000  
 cctgaccgga aacgccgcga gtgtcctggc cggccggctt tcctacacgt tcggccttga 30060  
 ggggcctcgc gtggcggttg ataccgcgtg ttctcgtcgc ttggtggcgt tgcatttggc 30120  
 gtgccagtcg ctgcgttcgg gtgagtgtga tatggcgttg gcagggtgtg tgacgggtgat 30180  
 gtcgacaccc ctggctttcc tcgagttctc tcgtcagcgc ggtttggcgc cagatggtcg 30240  
 gtgcaagtgc tttgcggccg ctgcggatgg caccgggttg ggtgagggtg ccggcctggt 30300  
 gttgctggag cggttgtcgg atgctcgtcg gaatggtcac cgggtgttg ccgtggttcg 30360  
 cgggtctcgc gtgaatcagg atggtgcgtc gaatggcctg actgcgccga atggtccgtc 30420  
 gcagcagcgg gtgattcggc aggcctcgc gaatgcgggg ctgtcggcgt ccgatgtgga 30480  
 tgtcgtggag gcgcacggga ccggtaccgg gctcggggat ccgatcgagg cgcaggcgt 30540  
 gatcgcgaca tatgggcagg agcgggatcc tgagcggggc ctgtggctgg ggtcgatcaa 30600  
 gtccaacatc ggccacacgc aggcggcggc cgggtgtggc ggggtcatca agatggtgca 30660  
 ggccatgcgg cacggggagt tgctgcgac gttgcacgtg gacaagccca ctccacaggt 30720  
 ggactggtct gccggggccg ttcggtcctt caccgggaac acgcccctgg ccgagagcgg 30780  
 ccgtcctcgt cgagcggggg tgtcgtcgtt cgggatcagc ggcaccaacg cacacctcat 30840  
 cctcgaacaa ccaccgtcgg aaccagcgga gatcgaccaa tcggatcggc gggtaactgc 30900  
 gcatccagcg gtgatccgtt ggatgtgtc ggctaggagt ctgcagcgc tcgaggccca 30960  
 agcggctcgc ctgcaggccc ggctggaccg gggctcctggc gcttctcgc tggatttggg 31020  
 gtattcactc gcgaccactc gttctgtgct ggacgaacgc gccgtcgtgt ggggtgccga 31080  
 tcgggaggca ctgctgtcca ggctggcagc gctcgcgat gcccgacgg cgcgggggt 31140  
 gataacgggc tctgogaatt ccggtggccg catcggatc gttttttccg gtcagggcag 31200  
 tcagtggctg gggatgggaa aggcgttgtg cgcggtttc ccggcgttcg cggacgcctt 31260  
 cgagggaagc tgcgacgcgc taagcgcaca cctgggcgcg gacgttcggg gtgtgctgtt 31320  
 cgggtctgat gagcagatgc tcgaccggac gctgtgggcg cagtcgggga tcttcgcgg 31380  
 tcaagtccgc ctctgggat tgcgtaggtc gtggggcgtg cggccggccg cgggtctggg 31440  
 gcaactcgtc ggcgagttgg ctgcggcga cgcggctggt gtgttgcct tgccggacgc 31500  
 tgcacggttg gttgcggctc ggcccacct gatgcaggca ttgccaccg gcggcgcaat 31560  
 gctcgcggtc gccaccagc aggcggcgg cggaccgctg ctttcggggg tgtgcgatcg 31620  
 ggtcagcatc gctgcgatca acggcccga gtcggtagtg ctctcgggcg accgcgatgt 31680



gctcgtggag ctgcgaggcg aattcgatgc ccgagggcctt aggaccaaat ggttgccgggt 31740  
ctcccatgct ttccactcgc accggatgga accgattctg gacgagtacg cggaaaccgc 31800  
caggtgcgtc gagttcgggtg aaccgggtggt gccgatcgtc tccgccgcga ccggtgcgct 31860  
ggacaccacc ggactgatgt gcgcggccga ctactggacg cgccaagtgc gtgatcctgt 31920  
ccgcttcgga gacggtgtcc gggcgctcgt cggccaaggc gtggacacga tcgtcgagtt 31980  
cggcccggac ggggcgttgt cggccctggt cgagcagtgc ttggccgggt ccgaccaggc 32040  
tgggaggggtg gcggcgatcc cgctgatgcg cagggaccgc gatgaggtcg agaccgcggt 32100  
ggcgccctg gcgcacgtgc acgtccgcgg tgggtgcggtg gactggtcgg cttgcttcgc 32160  
cggcaccggc gcccgacccg tcgagttgcc cacctacgcc ttccaacgcc agcggtagctg 32220  
gctggccggg caagcggacg ggccgcggcg cgatgtggtt gccgaccgcg tcgacgcgcg 32280  
cttctgggag ttggtcgagc gcgccgatcc ggaaccgttg gtggatgaac tctgcatcga 32340  
ccgggaccag cccttcgggg aggtgctgcc cgttctggct tcctggcgcg agaaacaacg 32400  
ccaggaggcc ctgcgggatt cctggcgcta ccagggtgcgc tggagggtccg tcgagggtgcc 32460  
gtccgcagcc gccctccggg gcgtgtggct ggtggtgctt ccagctgacg tgccccgaga 32520  
tcaaccggcg gtgcgtcatcg acgcgtgat cgccgcgcgc gccgaggtcg cggtcctgga 32580  
attgaccgag caggacctcc aacgcagtgc gcttgtggac aagggtgcgcg ccgtcattgc 32640  
ggaccgcacc gaggtgacgg gtgtgttgct tctgttggcg atggacggca tgccctgcgc 32700  
ggcgcatccg cacctgtccc gtggtgtcgc cgctaccgtg atcctgacgc aggtgttggg 32760  
cgatgcgggt gtttcgcgcc cgctgtggct ggccacgacc ggtggcgctg aggcggggac 32820  
cgaggacggt ccggccgatc cggaccacgg cttgatctgg gggctcggca gggtcgtcgg 32880  
ccttgaacat ccgcagtggg ggggtggcct gatcgacctt ccggagacac tggacgagac 32940  
gtcccgaac gggttggtgg ccgcactcgc cgggacggcg gccgaagatc agctcgccgt 33000  
gcgttcatcc gggttgttcg ttcgcagagt ggtgcgcgca gcgcggaacc cccggtcaga 33060  
gacatggcgt agccggggaa cggtcctcat cacgggcgga acaggcgcgc tcggtgccga 33120  
ggtcgacga tggctggccc ggccgggagc tgagcacctg gtgttgatca gtccgcgcg 33180  
cccgaagct cccggcgag cgacctaagg ggccgagctg actgaactcg gcgtgaaagt 33240  
cacagtcttg gcctgcgatg tgacggaccg cgacgagctg gcggcggtgc tggcgccgt 33300  
tcccacggag tatccgctgt cggcggtcgt gcacaccgcc ggctcggga cgcctgcgaa 33360  
cctggccgag acgaccttg cgagttcgc cgacgtgttg tcggccaagg tcgtcggcg 33420  
ggcgaacctg gaccggctgc ttggcgggca accgttggac gccttcgtgc tgttctctc 33480  
gatctcggga gtttggggag ccggcgccca aggagcctat tcggccgcca atgcgtatct 33540  
cgatgccctt gccgagcgcc gacgggcttg cgggcggccg gcgacgtgca tcgcctgggg 33600  
tccgtggcg ggtgcgggca tggccgttca ggaaggtaac gaggcgcatc tccgcgaag 33660  
gggctggtta ccgatggaac cgagtcggc cctcttcgcg ctgcaacagg cctgtccca 33720  
acgagaaacc gccatcaccg tcgcagatgt ggactgggag cgattcgccg cctctttcac 33780  
cgcgccccc ccgcgaccac tgttggaaga gatcgtggat ctacggccc acaccgagac 33840  
cgaggagaag cacggtgccg gcgagctggg gcagcagctg gccgcactgc cgcgcgtga 33900  
gcgcggacac ctgctgctgg aggtggtgct ggcgaaacc gccagcacc tggggcacga 33960  
ttcgcgggag gctgtgcaac ccgatcggac cttcgccgaa ctgggcttcg attcgtgac 34020  
cgcggtagag ctgcgcaaca ggtgaacgc ggtgaccggg cttcgccctgc cgcgacgt 34080  
ggttttcgac caccgcagc cgctggcgtt gtccgaacag ttggttcgg cctggtcgc 34140  
ggagccggac aacggcatcg aatcgctgct cgccgagctc gacaggctgg ataccagtt 34200  
ggcgcaagg ccttcgatcc cactggaaga ccaggccaag gtggcggagc gcttgacgc 34260  
actcctccc aagtgggacg gggcgctga cggcacggcc agagcgacgt caccccaatc 34320  
gctgacggcg gccacggacg acgaaatctt cgacctcatc gaccggaagt tccggcgctg 34380  
accgccctt cctgcctca gctcccctga ttactggaac ggtgtatttc gatggccaat 34440  
gaagaaaagc tccgcgagta cctcaagcgt gtcgtcgtcg aactggaaga ggcgacgaa 34500  
cgctgcacg agttggagcg ccaggagcac gaccccatcg cgatcgtgtc gatgggatgt 34560

cggtatcccg gtggcgctcc cactccggag gagctgtggc gactggctgt cgacggagga 34620  
 gacgcgatcg cgaacttccc cgaagaccgt ggctggaatc tggacgagct gttcgatcct 34680  
 gatccgggccc gagccgggac ctccctacgtc cgcgaggggtg gtttcctgcg cggggctcgcg 34740  
 gacttcgatg ccgggctctt cgggatcagt ccgcgcgagg cacaggcgat ggacccgcaa 34800  
 cagcgggttg tgctggagat ctctgtgggag gtgttcgagc gcgccggcat tgaccggttt 34860  
 tctttgcggg gtaccaagac cgggtgtgttc gcgggcctga tctaccacga ctacgcgtcg 34920  
 cggtttcgca agacccccgc ggagttcgag ggttacttcg ccaccggcaa cgcgggcagc 34980  
 gtcgcacccg gccgggtggc ttacaccttc gggtagagg gcccggcggt caccgtggac 35040  
 accgcctgct cgtcgtccct ggtggcgctg cacttgccct gccagtcctt gcggctgggc 35100  
 gaatgcgacc tggccctggc cgggtggcatt tcgggtgatg ccacgcggg agccttcgtc 35160  
 gaggtcagcc ggcaacgcgc actcgccctg gatggccggt gcaagccctt cgcggatgcc 35220  
 gccgacggca ccggctgggg cgagggcgcc ggaatgtgc tgctggaacg gctgtcggac 35280  
 gcacgacgaa acggccaccc ggtgctggcg gcgggtgtcg gttccgcgat caaccaggac 35340  
 gggacgtcca acggcctgac cgcgcccgag ggtcccgcac agcagcgagt gatccgcaa 35400  
 gccctggcga acgcggggtt gtgcgccgcc gaggtcgatg tggtcgaggc gcacggcacg 35460  
 ggcacggcct tgggcgaccc gatcgaggcg caggccctga tcgccaccta cggggcgaa 35520  
 cggtcggcgg atcatccgt gctgctgggt tccctcaagt cgaacatcgg ccacacccag 35580  
 gctgccgccc gtgtggccgg ggtgatcaag tcggtcctgg ccatcaggca ccgggagatg 35640  
 ccccgacgcc tgcacatcga ccagccatcg cagcacgtgg actggtcggc gggcgcggtg 35700  
 cggtgctca cggacagcgt tgactggccg gatctcgga gcccgcgccg agcaggggtg 35760  
 tccctgctcg gcatgagcgg taccaacgca cacctgatcg tcgaggaagt atccgacgag 35820  
 ccggtctcgg gcagtaccga gccgaccggg gcatttccct ggccgctgtc cggcaagacg 35880  
 gagacggcat tgcgcgagca ggctgccgag ttgctctccg tagtgaccga gcaccggag 35940  
 ccgggactgg gggacgtcgg gtactcgctg gccaccggtc gcgctgcgat ggagcaccgg 36000  
 gctgtcgtgg ttgccgacga tcgggactct ttcgtcgccg gactgacggc gttggctgcg 36060  
 ggcgttccgg cagccaacgt ggtgcagggc gcggccgact gcaagggaaa ggtcgcgttc 36120  
 gtgttccccg gccagggtc gcattggcag gggatggcga gggaactgtc cgaatcctcg 36180  
 ccggtgttcc ggcggaagct ggcggaatgc gcggcggtta cgtggactgg 36240  
 tcgctgctcg gcgtccttcg cgggtgatcc gatgcaccg cgtggatcg cgacgacgtg 36300  
 attcagctcg cgtgttccg catgatggtg tcgctggccg aactgtggcg ttcgtgcgga 36360  
 gtggagccc cgcgggtggt cggtcattcc caggcgaga tcgccgcgc ccatgtggca 36420  
 ggcgctttgt ccttgactga tgcggtgcgc atcatcgtg cccgctgca tcgggtgtcg 36480  
 gcgctgaccg ggaaggaggg catgctcgc attgccttc cggaaagcgc ggtggtgaag 36540  
 cgaatcgag gctgcgga gctgaccgt gcggcggtca acggaccgg ctccactgtc 36600  
 gtttccggcg aaccgtcggc tctggagcgt ctgcagaccg aactgaccgc ggaaaacgtg 36660  
 cagaccggc ggggtggaat tgattacgcc tcgcatcgc cgcagatcg gcaggtccag 36720  
 ggccggcttc tggaccggt gggcgaagtc gggtcgaac ctgctgagat cgctttctac 36780  
 tcgacggtca ccggcgagcg gacggacacc ggccgactcg acgccgacta ctggtaccag 36840  
 aaccttcggc agccgctccg cttccagcag accgtcgccc ggatggcaga tcagggtat 36900  
 cggttcttcg tcgaggtgag ccgcacccg ctgctcaccg ccggaatcca ggaaacgctg 36960  
 gaagccgcgg acgcggcg ggtggtggtc ggttcgctgc ggcgtggcga gggcggtcc 37020  
 cggcgctggc tgacttcgct ggccgagtgc cagggtgcgc gactgcgggt gaattgggaa 37080  
 caggattcc tcaacaccgg agcccgacgc gtgccgctgc cgacctacc gttccagcgg 37140  
 cagcgggtact ggttgagtc cgcgagtac gacgcggcg atctcggttc ggtgggcttg 37200  
 ctctccgccc agcatccct gctcggggt gcgggtgacgc tggccgatgc gggcggttc 37260  
 ctgctgacc gcaagctgtc ggtcaagacc cagccctggt tggccgacca cgtggtcggc 37320  
 gggcgcatcc tgetgcccgg caccgcgttc gtggaaatgc tgatacgcgc cgcggaccag 37380  
 gtcgggtgcg atctgatcga ggagttgtcc ctgacgactc cgtggtttt gcccgcgacc 37440

ggtgcggtgc	aggtgcagat	cgcggttggc	ggtccggacg	aggccggggc	ccgctcggtc	37500
cgcgtgcatt	cctgtcgaga	cgacgccgtg	ccgcaggact	cgtggacctg	ccacgcgacc	37560
ggcacgttga	cctccagcga	tcaccaggac	gccggccagg	gccccgatgg	gatttgggccg	37620
cccaacgatg	ctgtcgcggt	tccgctggac	agcttctacg	cccgcgcagc	tgagcggggc	37680
ttcgatttcg	gcccggcggt	ccaggggttg	caggcggtt	ggaagcgcg	agacgagatc	37740
ttcgccgagg	tcggcctgcc	caccgcacac	cgcgaagacg	ccggcagggt	cggaatccac	37800
cctgctctgc	tgatgcggc	actgcaggcg	ctgggcgcag	ccgaagagga	tccggacgag	37860
ggatggctcc	cgttcgcgtg	gcaagggtgt	tccctcaaag	cgacggggcg	actttccctt	37920
cgggtgcacc	tcgttccggc	gggcgcgaat	gcggtgtcgg	tggtcacgac	cgacacgact	37980
ggccaagccg	tgctctccat	cgattcgctg	gtgctgcgcc	agatttcgga	caagcagttg	38040
gcagcggccc	gtgcgatgga	acacgagtc	ctgttccggg	tcgactggaa	gcgaatctcg	38100
cccgccgctg	ccaagccggt	ctcctgggca	gtgatcggca	atgacgaact	cgcccagacc	38160
tgccgctcgg	cacttggcac	ggaactccac	cccgaacctga	ccgggttggc	tgacccgccc	38220
ccggacgtcg	tggtggtgcc	atgcggtgcg	tctcgccagg	acttggaagt	tgcttccgag	38280
gcacgtgccg	cgacacaacg	catgcttgac	ctgatccagg	attggttggc	ggcggcgcg	38340
ttcgccggat	ctgcctggt	ggttgtgacg	tggtgtcgg	cgtcgacagg	tcccgccgag	38400
ggtgtttccg	acctggtgca	tgctgcgtcg	tggggttgt	tgcttcggc	gcagtcggag	38460
aaccggacc	gattcgtgtt	ggtcgatgtg	gacggaaccg	ccgaatcatg	gcgtgcgtc	38520
gcggcggccg	tgcttccgg	agaaccgcag	ctggcggtgc	gcgccggtga	agtcggggtg	38580
cctgcctcgg	cgcatgtgt	tgccgccgag	gacagccgga	tcccagtgcc	cggtgcggat	38640
gggacggtgt	tgatttccgg	cggtagcggc	ctgctggggc	ggttggttgc	ccggcatttg	38700
gtggcggagc	gcggtgtccg	cgcctggtg	ctcgccgggc	gacgcggctg	gagcgcccc	38760
ggggtcaccg	acctggtgga	tgagttggtg	ggcctgggag	ctgcggtcga	ggtggcgagc	38820
tgcatgtcgg	gggatcgggc	ccagttggac	cggctgctga	cgacgatctc	ggcagagttc	38880
ccgctgcgcg	gagtgggtgca	tgccggccgg	gcacttgccg	acggggtcgt	cgagtcgctg	38940
acaccagagc	acgtggcaaa	ggtgttcggc	ccgaaggccg	ccggtgcgtg	gcacctgcac	39000
gagttgactc	ttgatctgga	tctctcgctt	tctcctcgtt	ctccggcgtg	39060	
gcgggggctg	cgggtcaggg	aaactacgcg	gcggcgaacg	cgttcctgga	cggcctggct	39120
cagcaccggc	ggacggcggg	gctgcctgcg	gtgtcgctgg	cttggggctt	gtgggagcag	39180
cccagcggga	tgaccggagc	gctcgatgcg	gcgggccgta	gccgcattgc	gcgcaccaat	39240
ccgccgatgt	ccgcgccgga	cgggttgcgg	ctgttcgaga	tgccgtttcg	cgttccgggc	39300
gaatcgcttc	tggttccggt	ccacgtcgac	ctgaacgccc	tgccgcgtga	tgccggccgac	39360
ggcgggtgtg	ctgcgttgtt	gcgcgacctg	gtgccagcgc	ccgtgcggcg	gagcgcggtc	39420
aacgagtcgg	cggacgtcaa	cggctctggt	ggtcggctgc	ggaggctgcc	ggacctggat	39480
caggaaaccc	agctgttggg	tttgggtgcg	gagcatgttt	cggcggtgct	ggggcattcg	39540
ggtgcggtcg	aggtcggggc	cgatcggtgt	ttccgggatt	tggtttttga	ttcgttgtcc	39600
ggtgtggagt	ttcggaaccg	gcttggcggg	gtgctggggc	ttcggttgcc	ggctactgcg	39660
gtgttcgact	atccgacacc	gcgggcggtg	gttcggttct	tgctcgacaa	actgattggt	39720
ggcgtggagg	ctccgactcc	cgcaccggcg	gctgtggcgg	cggtgactgc	tgacgatccc	39780
gttgtgatcg	tggggatggg	ctgtcgttat	ccgggtgggg	tgctcctcgc	ggaggagctt	39840
tggcggtttg	tggccggggg	cttggatgcg	gtggcggagt	tcccggaacg	tcgtggctgg	39900
gatcaggcgg	ggttgttcga	tccggatccc	gatcgtcctg	ggacctcgta	tggtgtgtgag	39960
ggtggcttcc	tgcgagatgc	ggcagagttc	gatgccggtt	tcttcgggat	ttccccgcgt	40020
gaggcggttg	cgatggatcc	gcagcagcgg	ttgctgctgg	aagtcgcttg	ggaaaccgtg	40080
gagcggggcg	ggattgatcc	gctttcgttg	cgggggagcc	ggaccggcgt	gttcgcgggg	40140
ctgatgcacc	acgactacgg	cgcgcggttc	atcacgaggg	cgcgggaggg	tttcgagggg	40200
tatctaggta	atggcagcgc	gggaggcgtg	ttttcgggtc	gggttgcgta	ttcgtttggt	40260
ttcgaggggc	ctgcggtgac	ggtggatacg	gcgtgttcgt	cgtcgttggt	ggcgctgcac	40320

ctggcgggtc aagcactgctg gtctggtgag tgtgatctgg ctcttgctggg tgggtgtgacg 40380  
 gtgatggcca cgccggggat gttcgtggag ttttcgcgtc aacggggctt ggcggcggat 40440  
 gggcgggtgca agtcgtttgc ggcggctgctg gatggcaccg gttggggaga aggcgcgggc 40500  
 ttggtgttgt tggagcggct gtcggatgcc cggcgcaacg ggcacgcggg tctggcggtc 40560  
 gtgcggggta gcgcggtgaa tcaggatggg gcgtcgaatg gtttgacggc gccgaatggg 40620  
 cctcgcagc agcgggtgat cacgcaggcg ttggcgagtg ctggtttgtc ggtgtctgat 40680  
 gtggacgccg tggaggcgca tgggactgga accaggcttg gtgatccgat tgaggcgcag 40740  
 gctctgattg ccacttacgg gcaggggagg gatagcgatc ggccgttgtg gttggggctg 40800  
 gtgaagtoga atattggtca tacgcaggcg gcggcgggtg tcgctggtgt gatcaagatg 40860  
 gtgatggcga tgcggcacgg gcagctgcc gcgacgttg atgtggatga acctacgtcg 40920  
 gaagtggatt ggtcggcggg ggatgtccag ctctcaccg agaacacccc ctggcccggc 40980  
 aacagccatc ctgcggcggg gggcgtgtcg tcgttcggga tcagcggcac caacgcacac 41040  
 gtcctctcg aacaagcctc gaaaacacca gacgagactg cggacaagag cgggtccgat 41100  
 tcggaatcga cgtggacct tccagcggtc ccgttgatcg tgcggggag aacaccggca 41160  
 gcgctcagcg ctccaggcgag cgcattgttg tcctatttgg gtgagcgtgg cgatatttcc 41220  
 acgctggatg cggcgttttc gttggcttcc tcccgggccc cgttgaggga gcgggcgggtg 41280  
 gtgctgggag cggaccgcga aacgttgttg tccgggttg aagcgtggc ttccggctcg 41340  
 gaggcttctg ggggtgtgtc gggatccccg gtctctggcg ggggttgggtt cgtgttcgcc 41400  
 ggtcagggcg gacagtgtt ggggatgggc cgggggctct actcgtttt tccggtgttc 41460  
 gctgacgcgt ttgacgaagc atgtgccgga ctggacgcgc atctggggca ggacgtgggg 41520  
 gtccgggatg tgggtgttgg ttccgacggg tccttgttgg atcggacgct gtgggcccag 41580  
 tcgggttgt tgcggttgca ggttggttt ctgagcctgc tgggttcgtg ggggtgccgg 41640  
 ccgggtgtgg tgctgggcca ttccgtccgc gagttccgg cggcggttgc ggcgggagt 41700  
 ttgtcgttgc cggatgcggc tcggatgggt gcgggtcgtg cccggttgat gcaggcgttg 41760  
 ccttctggcg gtgccatgtt ggcgggtggc gctgggtgag agcagctgcg gccgttgtt 41820  
 gccgatcggg ttgatgggtc gggatcgcc gcggtcaacg ctctgagtc ggtggtgctc 41880  
 tccggcgatc gggaggtgct tgacgacatc gccggcgcgc tggatgggca agggattcgg 41940  
 tggcggcggg tgcgggttcc gcatgcgttt cattcgtatc ggatggaccc gatgttgacg 42000  
 gagttccgcg aaatcgcacg cagcgtggac taccggcgtg gcgacctacc ggtcgtgtcg 42060  
 acgttgacgg gtgagctcga caccgcaggt gtgatggcta cgcggagta ttgggtgcgt 42120  
 caggttcgag agcccgtccg ctccgccgac ggcgtccggg tgctcgcgca gcaaggggtc 42180  
 gccacgatct tcgaactcgg ccctgatgcg acgctgtcgg ccctgattcc cgattgtcat 42240  
 tcgtgggctg atcaggccat gccgattccg atgctgcgta aagaccgtac ggaaaccgaa 42300  
 actgtggctg ccgcgggtgg gcgggcgcac acgctggtg ttccggctga atggtcggcg 42360  
 tatttcgcgg gcaccggggc acggcgggtc gagttgccga cgtatgcctt ccagcggcag 42420  
 cggtaactggc tggaaacatc ggattacggc gatgtgacgg gtatcggcct ggctgcggcg 42480  
 gagcatccgt tgctgggggc cgtggttgcg ctggccgatg gtgatggat ggtgctgacc 42540  
 ggccggttgt cgggtggggac gcatccgtgg ctggcccagc atcgcgtgct gggcgaggtc 42600  
 gtcgtccccg gcaccgccat cctggagatg gccctgcacg caggggcgcg tctcggctgt 42660  
 gaccgggtgg aagagctcac cctggaaaca ccgctggtgg tccccgaacg gcggcgggt 42720  
 gccggtagtc gtggccctgc gggagggacc acagtttcaa ttgaaactgc ggaagaacgt 42780  
 gtgcggacga acgacgcat cgaatccag ctgctggtga acgcacccga cgaaggcgg 42840  
 cggcgaaggg tgctcgtgta tccccgccg gccggtgggt cgagaggtgg ggggtggacg 42900  
 cgccacgcca ccggcgaact cgtcgtcggc accaccgggt gtagggcggt tctgatttg 42960  
 tcggctgagg gtgccgagtc gattgctctc gatgagttct acgtcgtctt ggccggaaac 43020  
 gggttcgagt acgggccgtt gttccagggg cttcaggcgg catggcgtcg tggtgacgag 43080  
 gttctcgcgg aaatcgcgcc gccggccgag gccgatgcga tggcgtcggg atacctgctc 43140  
 gaccacgcgt tgctggatgc cgcgtgcag gcgtccgcgc tcggcgaccg cccggagcaa 43200

ggcggcgcgt ggctgccgtt ctcattcacc ggcgtcgaac tttccgctcc ggcagggacg 43260  
 atcagcaggg tgcggctgga gaccaggcga cccgacgcga tatcggtggc cgtgatggat 43320  
 gagagtgggc ggttgctcgc ctcgatcgat tctctcaggc tacgaagcgt gtcgtcggga 43380  
 cagctggcga atcgggacgc tgtccgcgac gcgctgttcg aggtgacctg ggagccggtg 43440  
 gcgacgcagt cgacggaacc gggtcgctgg gccctgcttg gtgatactgc ctgcggtaaa 43500  
 gacgatctca tcaaaactcg aacggattcc gccgaccgct gcgcggatct ggcggcgcta 43560  
 gccgagaaac ttgattccag cgcgctgggt cctgatgtcg tggctctactg cgccggagaa 43620  
 caggcggatc ccggcaccgg cgcagccgca cttgcggaga cccagcagac gttggctctg 43680  
 ctccaagcgt ggttggtgga gccgcgggtg gccgaggcac gtctgggtgg ggtgacgtgt 43740  
 gcagcgggtg cgacggctcc gagtgcgggt gcatcagagc tggcacatgc gccgttgttg 43800  
 gggttgttgc gtgccgcgca ggtggagaac ccggggcagt ttgtgctggc ggacgtcgac 43860  
 ggaaccgccg aatcgtggcg tgcgttgccg agtgcggttg gctcgatgga accgcagttg 43920  
 gccctgcgga agggcgcggt gcgagcgcgc cgcttggtt cggtcgcccg gcagatcgac 43980  
 gtgcccgcgg ttgtggcgga tcccagaccga accgtgctga tttcgggcgg cacgggcctg 44040  
 ttggggggcg cggttgcccg ccacctgggt accgaacgcg gtgtccgcgg attggtgttg 44100  
 acgggcccgtc gtggctggga tgcctctgga atcaccgagt tgggtgggtga gctgaacggc 44160  
 ctcggtgccg tggtcgacgt ggtggcgtgc gacgtcgcgg atcgtgctga tctggagtcg 44220  
 ttgtggtggc cggttcccgg ggaatttccg ttgtgcggcg tgggtgcatgc cgcggggggcg 44280  
 ctggccgacg gggatgatga gtcgttgtca ccggacgacg tgggagcggg gttcggcccg 44340  
 aaggcggcg gggcgtggaa tctgcacgag ctgactcgtg atacggacct gtcgttcttc 44400  
 gcgttgttct cctcgttttc cgggtgttgc ggcgtcctg gtcagggcaa ttatgcggcg 44460  
 gcgaacgcgt tccctggacgc attggcgcgt taccggcggt cacagggact gcctgcgggtg 44520  
 tcgctggcct ggggcctgtg ggagcagccg agcgggatga cggagacgct cagcgaggtc 44580  
 gaccggagca ggatcgcgcg cgccaaccgg ccgttgtcca ccaaggaggg attgcggctg 44640  
 ttcgatgccg ggctggcgct ggaccgggca gcggtagttc cggcgaagtt ggacaggact 44700  
 ttctggccg agcaggcgcg gtcgggctcg ctgcccgcgt tgttgacggc actggtaccc 44760  
 cccatccgtc gtaataggcg ggctagcgga accgagctcg cggacgaggg caccctgctc 44820  
 ggggtggtgc gggagcatgc cgcggccgtg ctggggatt cgcgcgcggc tgacgtcggg 44880  
 gtcgagcgcg ctttccggga tctgggtttt gattcggtgt ctggtgtgga gttgcggaac 44940  
 cgccttgccg ggggtgctggg ggtgcggttg ccggcgactg cgggtgttcga ctatccgacg 45000  
 ccgagggcgc tggcccgggt cctgcaccag gaactggcag acgagatcgc tacgacgcca 45060  
 gcgcccgtga cgacgaccag ggcaccggtc gccgaagacg atctcgtcgc gatagtcggg 45120  
 atgggatgcc gttttcccg tcagggtgtcc tgcgcggagg agctctggcg tttggtggcc 45180  
 gggggcgtgg atgcggtcgc ggacttccca gccgatcgcg gctgggatct ggcaggcttg 45240  
 ttcgatccgg acccggaacg ggctgggaag acctacgtgc ggaaggggc cttcctcacc 45300  
 gacgccgatc ggttcgatgc gggtttcttc gggatttccc cgcgtgaggc gttggcgatg 45360  
 gatccgcagc aacggctgtt gctggagctg tccctgggagg ccattgaacg ggcagggatc 45420  
 gatccgggtt cgctgagggg gactcggacc ggtgtgttcg cggggctgat gtaccacgac 45480  
 tatggcgcgc ggttcgccag ccgagccccg gaaggtttcg aggggtatct cggcaatggc 45540  
 agtgctggga gtgtcgcgtc gggccggatt gcgtactcgt ttggtttcga gggctcctgcg 45600  
 gtgacggtgg atactgcgtg ttcgtcgtcg ttggtggcgt tgcatttggc gggtcagtcg 45660  
 ttgcgttccg gcgaatgcga tctgcacctt gccggtgggt tgacggtgat gtcgacgccc 45720  
 gggacgtttg tggaattctc ccgtcagcgg ggcctggcac cggacggggc gtgcaagtgc 45780  
 ttgcgaggga gcgcggacgg taccggttg ggtgaggggt ctggtttggt gttggtggag 45840  
 cggttgtcgg atgctcggcg gaatgggcat cgggtgttg cgggtggttcg tgggtcggcg 45900  
 gtgaatcagg atggtgcgtc gaatggcttg accgcgcgca atggtccctc gcagcagcgg 45960  
 gtcattccagc aggcgttggt gagtgccgggt ctgtcgggtg ccgatgtgga tgccgtggag 46020  
 gcgcatggga ccgggaccag gttgggtgat ccgattgagg cgcaggctct gattgctacg 46080

tatgggagcg atcgtg cggtaggcg ttgtggttgg ggtcg gaa gtccaacatc 46140  
 ggtcatacgc aggcggcggc ggggtgttgc ggtgtgatca agatggtgat ggcatgagg 46200  
 cacgggcaac ttccgcgcac gctgcacgtg gatgcacccct cctcgcagggt ggattggtcg 46260  
 gcggggaggg tccagctcct gacggagaac acgcccctggc ccgacagtgg tcgcccctgt 46320  
 cgggtggggg tgctgctcgt cgggatcagc ggcaccaacg cgcacgtcat cctggaacag 46380  
 tccacggggc agatggatca ggcagcggag ccgatttoga gtctgttct ggatgttccg 46440  
 gtggtgccgt ggggtgtgtc gggcaaaaca cccgaagcgc tatccgcca ggcggcaacg 46500  
 ttggcgacct atttggacca aaatgttgat gtctccctc tggacgttgg gatttcgctt 46560  
 gcggtgacct gttcggcgct ggatgagcgg gcggtgtgtc tggggtcggg tcgtgacacg 46620  
 ttgttgtctg gcctgaatgc gctggctgcc ggtcatgagg ctgctggcgt ggttacggga 46680  
 cctgtcggga ttggtggcgg gaccgggttt gtgttcggc gtcaaggcgg tcagtgggtt 46740  
 gggatggggc gccggttgta ctcggagttt ccggcgttcg ccggtgcttt cgacgaagca 46800  
 tgcgccgagc tcgatgcgaa cctggggagg gaagtcgggg ttcgggatgt ggtgttcggc 46860  
 tccgacgagt ccttgcctga tcggactttg tgggcgcagt cgggtttgtt cgcgttgacg 46920  
 gtcggtctct gggaattgtt gggtagctgg ggtgttcggc ccagcgtagt gctggggcat 46980  
 tcggtcgggg agctagccgc ggcgttcgcc gcaggtgtgc tgtcgatggc ggaggcggct 47040  
 cggctggtgg cgggtcgtgc gcggttgatg caggcgttgc cttctggcgg tgccatgctg 47100  
 gcggtgtccg cgaccgaggc ccgagtcggc ccgctgctcg atggggtgcg ggatcgtgtt 47160  
 ggtgtcgcag cggttaacgc tccggggtcg gtggtgcttt ccggtgaccg ggatgtgctc 47220  
 gatggcattg ccggtcggct ggacgggcaa ggtatccggt cgaggtggtt gcgggtttcg 47280  
 cacgcgtttc attcgcacgc gatggatccg atgctggcgg agttcgccga gctcgcacgg 47340  
 agcgtggact accggtctcc accgctgccg attgtctcga cgtgaccgg aaacctcgat 47400  
 gacgtggggc tgatggctac gccggagtat tgggtgcgcc aggtgcgaga gcccgccgc 47460  
 ttcgccgacg gtgtccaggc gcttgtggac caaggcgtcg acacgattgt ggaactcgg 47520  
 ccggacgggg cgttgtcgag cttggttcaa gagtgtgtgg cggagtcgg gcgggcgacg 47580  
 gggattccgt tgggtcggag agaccgtgat gagtcgcaa cgggtgctga cgttttggcg 47640  
 cagaccaca ctcgtggtgg ccggtggac tgggggtcat ttttcgctgg tacgaggga 47700  
 acgcaagtcg acctccac gtatgcctt caacgacgc ggtactggct ggagccatcg 47760  
 gattccggtg atgtgaccgg tgttggcctg accggggcgg agcatccgct gttgggtgcc 47820  
 gtggtgccgg tcgcgggcgg cgatgaggtg ctgctgaccg gcaggctgtc ggtggggacg 47880  
 catccgtggc tggcggaaca ccgctgctg ggcgaagtcg tcgtccccgg caccgcgtt 47940  
 ctggagatgg cgtggcgggc cggtagccag gtcggttgtg aacgtgtgga ggagctcacc 48000  
 ttggaggcac cgctggtcct gccggagcgg ggcgctgcgg cgggtcagtt gcgggtggg 48060  
 gctccggatg aggcggccg gcgcagttt cagctctatt ccgagggcg tgatgaagac 48120  
 ggcgactggc ggcggattgc ctccgggctg ttggcccagg ccaatgcgg gcccggcg 48180  
 gattcgacgg catggccgcc ggacggcgcc gggcaggtcg atctggcgga gttctacgag 48240  
 cgctcggcg agcgcggctt gacctacgt ccggtattcc aagggtccg cgcgcgatg 48300  
 cggcacggcg acgatattt cgcgaattg gccgggtcac cagacgcctc gggtttcggc 48360  
 atccaccogg cgctgctgga cgctgactg cagcgatgg cgcttgggtg ttcgcccga 48420  
 tcggaagcgc gtctgccgtt ttctggcgt ggcgccagc tgtaccgcgc tgaaggagca 48480  
 gcgcttcggg tacggtctc gccgctggg tccggtgcag tctcattgac gttggtggat 48540  
 gccacagggc gaogagtcgc tgcggtgaa tcgctttoga cgcgaccgg ctccaccga 48600  
 cagatcgggt ccggtcgcgg cgatcaagag cggctgctgc acgtcgagt ggtaaggctg 48660  
 gctgaatctg cggggatgtc tctgacctc tgcgcgttg tcggtttggg cgaaccggag 48720  
 tggcacgctg cgctgaagac cactggtgtc caagtcgagt cccatgcgga ccttgcttcg 48780  
 ttggccaccg aggttgccaa gcggggttca gctcctggtg cggtcacgt cccgtgccc 48840  
 cgaccccgag cgatgcagga gctgcgcacc gccgcgcgaa gggcgacgca acaggcgatg 48900  
 gcgatgctgc agcaatggct tgccgatgac cggttcgtca gtacggcct gatcctgctg 48960

acgcacgagg cggtctccgc agttgctgga gaagacgtgc tcgacctggt acacgcgcgc 49020  
 ctgtggggct tggctccgag cgcgcaagcg gagcaccgag accgattcgc cttgatcgat 49080  
 atggacgacg agcgagcatc gcagacggca ctcgccgaag cgctgactgc gggagaagcg 49140  
 cagctcgcgg tgcggtcggg agttgtgctg gcgccccgcc tcggccaggt gaaggtgagt 49200  
 ggaggtgaag cgttcaggtg ggatgaaggc accgtgctgg tcaccggcgg aaccggcggg 49260  
 ctcggggccc tgctcgcacg ccatctgggt agcgcccacg gtgtgcggca cctgttgctc 49320  
 gcaagtcgcc gtggtctggc ggcgcccgga gcgcatgagc tgggtggcga gctggagcag 49380  
 gccggcgccg acgtcgcggt cgtcgcgtgc gactcggcag atcgggactc gcttgccgag 49440  
 ctggtggcgt cggtgcctgc ggaaaacccg ttgcgggtgg tgggtgcacgc cgccggtgtg 49500  
 ctggatgacg gtgtgctgat gtcgatgtcg ccggagcgct tggacgcggt gttgcggccc 49560  
 aaagtggatg ccgcgtggta cctgcacgag ctgactcggg aactcgggtc gtcggcgctc 49620  
 gtgttgcttc cctcggctgc gggcctgttc ggcggtgcgg ggcagagcaa ttacgctgcc 49680  
 ggcaacgctt tcttgatgac cttggcgcat tgccggcagg cccaggggct gcccgcgctg 49740  
 tcgctggcct ccgggctgtg ggcgagatc gatggaatgg cggcgacact cgctgcggca 49800  
 gatgtggagc ggctgtcgcg ggcaggcatt ggcccgcttc cggcacccgg agggctggcc 49860  
 ttgttcgacg ctgccgttgg ctcggacgaa ccgttgctgg caccggtgcg actggatgtc 49920  
 gaagcactgc gtgtgcaggc ccgatccgtg cagaccgga ttccggaaat gctgcatggc 49980  
 atggcaatgg ggccaagccg 50000

<210> 3

<211> 25360

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 3

ccgcactccg ttcaactcca gggttgagcc gttgcacgaa cggctggccg gattgtcgga 60  
 gggcgaacgt cggcagcaag tgctccagcg cgtccgcgcc gatatcgcg tggactggg 120  
 gcacggcagg tcgagcgatg tggacatcga gaagcctttg gccgagctgg gtttcgactc 180  
 gctgacggcc atcgaactcc gcaaccgtct cgtaccgcc accggactgc ggttcccg 240  
 gacgtggcc ttcgaccacg gcaactgcggc ggcactcgcc cagcacgtgt gcgagcagct 300  
 aggcaccgag accgcgcggc caccgaggcg aaccgacgac aacgacgcca cggagccgct 360  
 gaggtcgctc ttccaacagg cgtatgcggc tggccggata cttgacggga tggatttgg 420  
 gaaggtcgct gccagttgc gaccggtgtt cggttcgcc ggcgagctgg aatccctgcc 480  
 gaaacccgct cagctttccc gtggtcccga agagcttgcc ttggtgtgca tgccggcgct 540  
 gatcgggatg ccgcccgac agcagtacgc gcgcatcgcc gccgggttcc gcgatgtgcg 600  
 ggacgtttcg gtgatccga tgctggatt cattgcggga gaaccgctgc cgtccgccat 660  
 cgaggtggcg gttcggacgc aggcggaggc ggtgctgcag gaattcgccg ggggctcgtt 720  
 cgtactggtc gggcattcct ccgggggctg gctggcgac gaggtagccg gtgagctgga 780  
 gcgtcgcggg gtcgtcccg ccggggctgt actgctggac acctacatcc ccggtgagat 840  
 cagcccgagg ttctccgtgg cgatggcca ccggacgtat gagaagctcg cgactttcac 900  
 ggacatgcag gatgtcggta tcaccgcgat gggcggttac ttccggatgt tcaccgagt 960  
 gactccgacg ccgatcgggt ctcgcacgct gttcgtgcgg accgaagatt gcgtcgcaga 1020  
 ccctgaaggg cggccgtgga cagatgactc ctggcgcca ggggtgactc tcgcgatgc 1080  
 caggtccag gtgcccggcg accatttctc gatgatggac gagcacgccg ggtccaccgc 1140  
 acaggcagtc gcgagttggc ttgacaaact caaccagcgc accgctcggc aacgctgacg 1200  
 ggcgtccttt taggacctc tggcgggcac cggccacccc ggcggtgccg ccttcggtg 1260  
 tccaggctcg ccgatcttga cggcgcacga tgcgcggcac gcgcgctgat cgtgattccg 1320



ctgccgctcg tggccatggg cctggcgaat catgtccttt cgggcaggc caaacgaatt 1380  
 cgcccgagcc cgcattccga ggtgaggggc acccttggtt ggctgagccg ctcaaggggtg 1440  
 cccctcacct cgaaattcgt ccgatttggg cgggtgacgc aaccccggtg ggcgtggtgc 1500  
 gtctttcttg ttgacagagc ggtgagaagc cgctgacaca cctgagagga aaaggggagc 1560  
 atgatgtca agcgccaccg tttgacgacc gccatcaccg gccttctggg gggagtactg 1620  
 ctggtcagcg gctgcggaac cgccgcgcga cttcagtcct cgccggcgcc cgggcatgac 1680  
 ggcgcgaatg ttggtatggc ctccggcggg ggccggcggg acatcggcac gtcgaactgc 1740  
 tcggaggccg atttcctcgc caccgcgaca ccggtgaaag gcgaccccg cagtttcate 1800  
 gtggcgtagc ggaaccggtc ggacaagacc tgcacgatca acggcgcggt gccgaacctc 1860  
 aagggcggtg acatgagcaa ctccgcgac gaggacctgc cggtcgagga cgtgcggcctt 1920  
 cccgacgcgc ccaaggaatt caccctccag cccggtcaga gcgcgtacgc cggcattggc 1980  
 atggtccttg ccgacagcgg cgacccgaac gcccatgtcc tcaccgggtt ccagtcctcg 2040  
 ctgccggaca tgtccgaggg ccagccggtc aacgtttctg gcgacggcaa cgtgaagtgc 2100  
 gccgcgaagt acctgcgagt cagctcgctg gtgtctaccg cagacgagct gcgctaaaac 2160  
 ccatgtgagt cccgcagatt cgacctcgcc gtgcggcgcc tccggcgaag cgtccgtacg 2220  
 tttgtcgttg tgaccagcgt tgttcacgtc cgggcgcagc gctggtacat actcaggcgt 2280  
 ctccggcgcc tccaacgggg cctggcatcc ggggcgcgtc agtgcggcgg cgtgacgcgc 2340  
 ttctctgtcg ggcgttgtca cgccgcgggc ctccaaccgg tcccgcgccg tcggagccgg 2400  
 tggtcacagc cgggtgtggc gcggccggag ccgacggtgc gcaccgcctg cccgagggcc 2460  
 tttttogaac cgacgaggac cagaccttc ttggcccggg tgaccgcctg gtagagcagg 2520  
 ttgcgctgca gcatcatcca ggccgttgtg gtcaagggga tcaccacgca cgggtattcg 2580  
 cttccctgcg aacgatggat ggtcaccgcg taggcgtgga ccagttcgtc gagttctgtg 2640  
 aagtcgtagt cgatgtcttc gtctcgtcg gtccgacgg tcattggtctg tgcttcgttg 2700  
 tcgagggcgg acacgacgcc ctgcgtgccg ttgaacacgc cgttggcgcc ctgtcgtag 2760  
 ttgttgcgga tctgcgtgac cttgtcgccg acgcggaaga tccgtccgcc gaaccgcgc 2820  
 tctggcaggg cctccctggc cggggtgatc gcttctgca acagctggtt cagcgcgcct 2880  
 gcacctgcgg ggcctcgatg catcggggcg aggacctgca cgtcgggtgc cgggttgaa 2940  
 cggaacttcc gcggaatccg gcgggcgacg acgtcgacgg tgagctcggc ggtcgggttcg 3000  
 ctttctcta cgtggaacag gaagaagtcg gtcagccctg gtgtcagcgg atagtccccg 3060  
 gcgttgattc ggtgcgcggt ggtcaccacc ccgactcgg cggcctgccg gaacacctcg 3120  
 ttgagccgca cgtgtggaat cggggtgcca ggggcgagca gatcgcgagc tacctcaccg 3180  
 gtcggaccg acgggagctg gtcgacgtcg ccgaccagca gcaggtgcgc gccgggcgcg 3240  
 atcgcttgg ccagtttgtt ggctaacagc aggtcgagca tggacgcctc gtcgaccacg 3300  
 acgaggtcgg cgtccagcgg gttgtcccg tctagggcg cgtccccgc cggctggagt 3360  
 tggagcaggg ggtgcacggt cgccgcgtcg tgtccggtga gctcgggtcag ccgcttcgcc 3420  
 gctcgtcccg tcggcgcggc gaggatcac ttggcctttt tcgctgagc taatgcgatg 3480  
 atcgaccgca cgggtgaagt cttgcgcgag cctggacctc cggtgagcac ggcgaccttc 3540  
 tcggtcaggg ccagcttgac ggcgcgctcc tgcgcctcgg cgagttcggc accggtagcg 3600  
 cggcgcaacc agtcgagggc cttgtgcca tgcagctcgg cgaagacggg catccggtcc 3660  
 gcgctggtgt tcagcagccg ggacagctgg ttggccagg cgacttcggc gcggtggaag 3720  
 ggcacgaggt agatcgcgac cgtcgccacc tcgtcgtcat cggtggggat ctctcgcg 3780  
 accacacctt cctcggtgac gagttcgggc aggcattcga tcaccagccc ggtgtcgacg 3840  
 gcgaggatct tcaccgcctc ggcgatcagc tcgttctccg gcaggtagca gttgccgtcg 3900  
 ccggtggact ccgacagcgt gaactgaagg ccgccttta cccgctgcgg ggagtgcgtg 3960  
 gggattccca ccgctttggc gatggtgtcg gcggtcttga aaccgattcc ccacacgtcg 4020  
 cctgccagcc ggtatggctc ttccttgacg gtccggatcg cgtcgtcgtg gtactgcttg 4080  
 tagatcttca ccgccagcga ggtcgagacg ccgacgcctt gcaggaagat catcacctcc 4140  
 ttgatcgctt tctgtcctc ccaagcgctc gcgatcagct tcgtccgctt cgggcgcgagc 4200



ttggggacct	cgatcagccg	cgcggggttc	tgctcgatga	cgteagagcgc	ggcgacgccg	4260
aagtggtcga	cgatcttctc	ggcgagtttg	gggccgatgc	ccttgatcag	gccagacccc	4320
aggtagcggc	ggataccttg	cacggtcgca	ggcagcacgg	tcgtgtagtc	gtcgacgtgg	4380
aactgccgcc	cgtactgggg	gtgcgacccc	caccggccgc	gcatgcgcaa	cgcctcgccg	4440
ggctgcgcgc	ccagcagcgc	gccgacgacc	gtcaccaggt	caccgccccg	gccggtgtcg	4500
atccgcgcga	cgggtgtagcc	gctctcctcg	ttggcgaaacg	tgatccgctc	cagcgtgccc	4560
tccagcaccg	cagtccacgt	ggccgactcc	cgtccttttt	ccaccgacaa	cacgtatcac	4620
gaacggctgt	caagcaaacc	ggcggtcacc	acatgcagcg	gcatctcccc	aacgcctcgg	4680
gctccggcgt	cagcgggtgg	gcgttcgcga	tgccttggtg	cggccgggtg	gagttgtaga	4740
ttttttcgtc	ctcgcgcagg	gcctggagta	ggtgccgctg	gctccagatc	aggggaaagc	4800
gccgggataa	accggcttga	cggaggagat	ggaagagctc	tacgtcgaat	ggccagcgac	4860
ccacgacgac	cccagatcat	gcgtcgacga	cctgcgaggc	cgtggcgaag	cgttgacagg	4920
gggcacgtgg	gggcgtgcga	tcaggtcgcg	gcgccggctc	agcgcggtgt	ccggacgcac	4980
cagcagccgc	gacggactgg	ttccgctcct	cgagcgcggt	gtcctcgatc	gtcttttttcg	5040
agctcgatgc	cggcgcgctt	cgcgagggtg	ccagtccagt	cggtcggcag	tcgtgggtcg	5100
agtctgcctt	ttgaacatct	ggtcaattca	gctttgaatg	atcgcttcggc	gagtattgtc	5160
ctgcccgcca	attgtttgct	cgccgtgcgc	cctgggacag	gagcgcggcg	ttcggggacg	5220
gcgtgctggc	gaaggtgtcg	attggctcgg	acggccggac	caggatcgcc	gacgtcgtga	5280
tcgccacgag	cgteggcgat	gcccggtact	atctggagca	ccagcgagtg	gacaacgact	5340
tccacgggcg	cgggtgcgctc	ctgatcatga	atcaagaact	gagaccgtca	gggtgaacgt	5400
gccgtcgtgg	ggcaactgtga	ggatgtgcat	cacacggccg	gggtaggtga	tcgaggaggc	5460
cctgatcgcg	ggctgccgcg	gctgtgccca	gtcgaatcgc	agcaccacct	ggccgcgggt	5520
gagcgtcagc	ccctgcgcgt	cgagtgcgac	tttgtcggag	tccttcacca	ccaacgggat	5580
ctgctcggtt	gccggtgcgt	tcgcctgcac	cgacctggtg	atccggtcgg	gccgcacggt	5640
cacatcggtg	acgaccgagc	cattcgcggt	gcggtagcgg	aaccggtatg	gctcgccgcc	5700
caggaactcc	gcctgctgcg	gtccgttcgc	gtccggcacc	tggcccggga	agaccgtcgc	5760
ccagcaggcg	tggtcgttgt	tgttcatcga	ctggacgacc	gtgcccgcga	ccgggtgcca	5820
caggaacgtc	agaccgggtg	ggcccagcga	ggtcgcgcgc	ctcgtgccga	agtagccgcc	5880
gaggtagaac	tgcggcctgc	gtaggaacag	gaagtcttgc	ccctggtcgc	tcgggatctc	5940
ggtgaagtgc	gcctgttggt	atggaagttt	cgcgatcgcg	gccgtgcgct	gggcgcgggt	6000
cggatagcgc	tcgcctgtaga	tcgcgtgggt	gaggatgcgc	ggtgaggtgt	cctgtttgac	6060
gagtgcgggc	accggtgccg	ggtcctcggc	ccaggccgct	cgtgcggcgg	cgaggtcttc	6120
gcggctggac	aggaacgcgc	ccaagttcgg	caccactggc	acgaactggt	tgttcagcgc	6180
ggtacgttcc	gggtccggtc	gcacgtcggc	gtagaaccgc	gccgaggtgc	gggagttggg	6240
ggcgacgttg	gcgaaccagc	cggagccgtc	gggttcgcgc	agcaggttgt	agctgagcca	6300
gtcgggtgtac	ttccgcgcga	tcgtgacgag	gttggcgttg	ttggactggc	gccacgcgtc	6360
ggccatctcc	ggcagcatca	cctcgaagtt	gtagttggtg	tcgcgcggcg	tcaactcgta	6420
gaagtagccc	gaaggactct	gcccgttcgc	ggcaaggcgc	gcgaacgcct	cgggtcaactg	6480
ctgccgcagc	gcggggtcgg	ggtcctgggt	cagcgcgagc	gaggagccgg	cgagcccggc	6540
ggtaacctgg	ttcgcgtagt	ggattgtggg	ctgccaggcg	ctaccgttgc	ccgggttgag	6600
gagccacgtc	atcgctttgc	gcagcgccga	ggtgatctgc	gactggcggt	gcggcaggat	6660
gttcgccgcg	cgcagcaggg	cgttggtctt	gctcaggtac	cccaaccgga	agccgggtggc	6720
ggcgagcccc	tgctcggtcg	gcgaatactc	cggccacgag	ccgtcatcgt	gctgcaaccc	6780
taggtagtgt	ccgagcccgg	cgtcgagcgc	ggcgagcaga	gtcgcgtcgc	cccgttacgg	6840
gttccacgtg	cgggattgcg	tggcgaaacca	ggcgagcgtg	tagacgtgct	cctggacgcg	6900
ggcgttgtag	gacaccgccg	ggctgcgcga	ccagcctccg	gcgaagaagc	cgctcgagtc	6960
catgtcggcg	accatcgggc	agaccgcggt	gaggtacgac	gcgaatcgct	gctcctcggg	7020
tgcgaacagg	cgccgggttcg	gtgtggtggg	cggcgggggc	ggcagtgcca	gtgcgcgggt	7080

cggcagtgcg gcgagcagtc cgagagcggc ggcgccggtc atcaggctgc gacggctgaa 7140  
 cgtagtccacg ggcctacctc cttgtggcgg atcaaccctc acccgctgcg tagccgcacg 7200  
 tcaagatgat aattcgaatt attatgggtg tgacgacgcg taggcccacg acgcagaatt 7260  
 cctgcccaatt cgtattggca agcgggggtg ctcgtggccc gacggctcac ccagcaggac 7320  
 atcgcccggga tggcaggagt cagccaggcc acggtgtcgc tgggtgctcaa caaccggaag 7380  
 gacggcaacg tccgcatcgc ggcggagacc cgtgcgcagg tactggaggt gattcgggaag 7440  
 accggctacg tcgcgaaccc gatcgcccgc aggatgcgcg atcggcaca cgcacatcctc 7500  
 ggcgtgttca cctacgaggc ggtgttcccc agcaccacg cgaacttcta ccagtgcgtc 7560  
 ctcgaaggca tcgaggaaca ggccgaggag gtcggctgtg acctgttgtt gttcaccagc 7620  
 gccaaggcca ccggggagcg gcggcggtatt ttcggcgacg acagccgggt gcggtctgcc 7680  
 gacggcagtc tgctgctcgg tcgcacggtc gaccgcgcag acctgaccca gctgctcggc 7740  
 gaaagcatcc cgtacgtctc catcgagcga cgcgacgacg cggggcggtcc ggtgccgcac 7800  
 gtcggggccg actaccgcac cgcggtgcga gacctggtg accgcgcggt cgcgctcggc 7860  
 caccgcgggt tcgcgtacgt ggggtctggt gggggcgcg agtcgtccgc ggatcggctg 7920  
 cgaggcttcc gcgaagccgt tgccgcacat ggcgtccaag ggatgcatgt ggagacccca 7980  
 cagctcgatc agctgcgcga agcgggcgtc accgctgtgc tcaccgaaga ggtgtcggac 8040  
 ggggccgcgc tcgtgctcgc cgggcgcgaa cgcgggctct ccgtgccggg cgacctcgcc-8100  
 gtgctctcgc tcggtgccgc taccgggtcg gcaccggacg acgacgtggc gctcaccggt 8160  
 ttccgcatcc ccaggcgcga gatggggcgc cgggcggtgc aggcgctgac cgaggtgctc 8220  
 gaaaacggca ccacaccgca agaactgctc ccgtgcgagt tcgtcgaggg ctcgacgctg 8280  
 ggcgacccac gcctttgacc aggaggaact gttgctcgac cacaccacgg acgttgctcg 8340  
 cgttggcggc ggactcggcg gtgtcgcgc cgcactcgcg ttgctgcgcg cgggccgccc 8400  
 ggtcgtgctc accgaggagt acgactggct cggcgccag ctgaccagcc aggcgctgcc 8460  
 gcccgacgag cacagctggg tggagcgctt cggcgtcacc gcgagctacc gggcgttacg 8520  
 cgacggcatc cgcgactact accgcgcga ctaccgctg accccgcgcg caccggcggtg 8580  
 gcgggagctc aaccccggtg cgggcaacgt gagcaggctc tgccacgagc cccgcgtcgc 8640  
 cgtcgcggtg atcgacgaga tgctggcgcc gttccgcggc agtggcaggc tgaccgtgct 8700  
 gcagccgtac cggccggtg ccgcgcacaa cgacggcgac cggatcgtgt cggtgaccgt 8760  
 tgcgacccgc gacaccggtg aacagatcga gctctccgc ccgtacatcc tggacgcgac 8820  
 ggagacgggt gaactgcttc cgttgtccag cacggagtac gtcaccggct tcgagtccac 8880  
 tctggacacc ggcgagccga gtgcgcccga cgtcgcgcag ccggcgaaca tgcaggcggt 8940  
 gtcggtgtgc ttcgtggctg accacgtcga cggcgaccac accatcgaca aaccggcgcg 9000  
 gtacgacttc tggcgcgct accagccgga cttctggggc gaccggatgc tgtcgttccg 9060  
 ctcccccaac ccgcgcacgc tcgcgatctc cgaacgtacg ttcacccga acccgagcga 9120  
 cgacccgctc ggcgtcgtgt cggaccagcg gctcagtgcg ggtgacagca atctgtggac 9180  
 gttccggcgc atcgccgcgc gtcgcaactt cgtcgagggt gcctacgaca gcgacatctg 9240  
 cctggtgaac tggccgatca tcgactactt cgagtgcgcg gtgatcgacg tgccggacgc 9300  
 cgacgcgcac atcgccgcgg cgcgggaact ctcacgttcg gtgctctact ggctacagac 9360  
 cgaggcgccg cgcacagacg gcggcaccgg cttccccggc ctccgcctgc gcggcgacgt 9420  
 caccggcagc gcggacggtc tcgcgcaggc gccgtacatc cgcgagtcca ggcgcacacg 9480  
 ggccgagcac acgatcgtcg aacaggacct ctcgctcgcg gtgcgcggcg acaagggtgc 9540  
 ggtgcagcac gccgacgccg tgggtgtcgg catgtaccgt atcgacctgc acccctccac 9600  
 cgggtggcgc aactacatcg acgtcgcgag ctgcccgttc gagatcccgc tcggcgcgct 9660  
 gatcccga cgggtggaga acctgctacc cgcgggcaag aacatcggca ccaccacat 9720  
 caccaacggt tcccaccggc tgcacccagt cgagtggaa gtcggcgagg tcgcgggcgc 9780  
 gtcgctgcc ttctgcctgg cgcaccgagt caccctcgc gcggtgcgca ataccctgg 9840  
 cctgctcgcg gacttccagc agtgtctgga acgcgacggg gtcgagctcc gctggccgga 9900  
 cgtgtccggc tactgacgca gggagacgaa aatgacaaag ctgtcacgac gactcacggc 9960

actcatcgtc	gcagggctgt	tcgccctcac	cggtcgcggt	ggtggatcaa	ccgcacagtc	10020
cggaccgaag	tcgtgçgca	tgaccgtgtg	gactgccaac	gcggcgcatc	tcaagctgct	10080
caacgacatc	gccgccgagt	acaaggcctc	gcacccggac	atcgccgaga	tcaagttcga	10140
ctcggtgccc	gccgacggct	acaccaccac	gctcaccacc	cagatcgccg	gcggtaacgc	10200
gccggacctg	gcctggatcc	tggaggagtc	ggcaccggac	ttcgtggcgt	ccggtgcgct	10260
cgccccgggtg	cgcggaaga	tcgagaaggc	cgacgagctc	gtgccgtccg	cgacgaagct	10320
gtgggagaag	gacggcgaac	tgtacgccta	cccgttctcc	acctcgccgt	tcggcttggt	10380
cgtcaacacc	gacctggtga	agggcgccctc	ggcggactgg	acctgggacc	aggcgatcgc	10440
ggctgcctct	gcgtcggcgg	ccgcctccgg	caagggcggc	ctggtactgc	cggacttcaa	10500
gtaccagaac	tgggcagtg	tgtcctctat	ctggcgcggc	tggggagctg	atgctgggag	10560
cgcgacgggt	cgctcggtgcg	ggttctccag	cagcgagatg	aacgacgcga	tgtccttctt	10620
gcacaaggcc	atcttcaccg	acaaggcgat	tccgggcccc	ggcacgacgg	tggacttctt	10680
cgccggcgac	gcggcgatgg	cgatcggcc	gatctcccgg	tccagtgcgt	tgaaggacgc	10740
gaagttcggc	tggacgctgc	tgccgctgcc	ggccggtcgg	aagggtgact	acgcggtgat	10800
cgggcaggcc	gggatcggtg	tgctgaagca	gtcgcaaac	gtcgacgccg	cgacggactt	10860
cctcgcttcc	ctcaccaacc	agaccaactc	cgccaagctc	gcgcagttct	tcccccggc	10920
gcggtcgctg	ctgctcaacg	cggagacgct	cgccaagagc	aaccggttga	tcaaggccga	10980
gcaactgcag	tccgtcgctg	tcgacggcat	caacaagggc	gtcgtgaagc	cgagccacaa	11040
gggtcaggag	gagctgaacc	agacgatccg	cgccgcgctc	gaccggtgt	ggaagccgga	11100
cgcgaacgtg	cagaacgtgc	tgaacgacgt	gtgcaccaag	atcaaaccgc	tgctggagaa	11160
caagtgcg	cggtcgca	ctccacgcgg	gccgtagccc	gcaaggggac	gtcgtactgg	11220
acgcagcgca	ggcgcgacaa	tctggttggt	tacctgttcg	tcgcgcccgc	gctgctcggc	11280
agcatcgctg	tcgtgctggg	gccgctggcc	ttggtcggct	ggtacagcct	caacgagtgg	11340
aacgtgctcg	ccggcacggt	cgagttcgctc	ggcgcgacga	actaccaaca	gctgctcgcc	11400
gacgagaagc	tgcgcgactc	gctggtggcg	accacttggt	tcgcgcgcgg	cctggtcggtg	11460
ctcaacctgt	cactagcgct	gctgctggcc	gtgctgctca	accagaagct	gagcggcacc	11520
acggtcttcc	gcacgttggt	cttctctccc	gtcctcggtg	cgctggtggc	gtggacgctg	11580
gtgtggcagc	tgatactgca	gccggagggg	agcgtcaacg	gactgctcgg	gttcttcggt	11640
gccgacgggc	cgaactggct	gcgcggtgag	tcgaccgcga	tggctcgggt	gatcgctggtg	11700
caggtgctca	agaacgtcgg	cctgaacatg	gtgctgttcc	tcgcggcgct	gcaaggcggtg	11760
ccgcagccgt	tgtacgaggc	ggcgaagatc	gacggggcg	gtgcctggac	ccggttccgc	11820
cgcatacct	tgccgttgat	cagcccgacg	atcctgctca	cgctgatcat	caccatcgtc	11880
ggctcgctgc	aggtgttcgc	gcagatcgcg	gtgctcacgc	agggcggtcc	gggcacgtcc	11940
acgaccgtac	tgatctacta	cctgtaccag	caggcggtcc	agttccacca	cttcggctac	12000
ggcgcgacga	tctcggtact	gctgttcgctc	atcgtcgcgg	cactcaccct	gctgcagtgg	12060
cagatgcgcc	ggaagtgggt	gctgcatgag	gcttaggggc	aagatcacgc	tatacgggct	12120
gctgtgcctg	ctgtgcgtgc	cttctggtgt	cccaacctgg	tggatgatca	cctcgctcgat	12180
gaagcccatc	agcgagatct	tctccacctc	gccgctgccg	tcggagtgg	cgttctccac	12240
ctaccggcag	gtgttcgaga	tgacgcggtt	cgcgacgacg	tactggaaca	gcctctacat	12300
cgcggtgac	gtcacgacag	gcacaatggc	ggtggcgcg	atggccggat	acgcgttcgc	12360
gcgtatccgg	ttcccgggcc	agaacgtgct	gttcgtgggc	gtcctgatcg	gcctgctcat	12420
cccgaagcag	gtcaccatcg	tgccgctggt	caagatgttc	cagtcgctcg	gcctgaocaa	12480
cacgcactgg	ccgctaata	tcgtgccgat	cttcggggcg	ccctgtgtgc	tggcgatctt	12540
catcatgcgc	cagttcttca	tcgccctgcc	gtcggagctg	gaggaggccg	cgcggatgga	12600
cggcctcgga	cgagcgggca	tcttctaccg	ggtggccctg	ccgctgtccc	ggcccgcgct	12660
gggcgcgctc	gcgatcttca	cgttcttgca	ctcgtggaac	ctctacctgg	agccgatcgt	12720
ctatctgtcc	acaccggaca	tgtacacggt	gccgcaagcg	ctcacgcagt	tcgtcgacaa	12780
ctacggcggc	ccgatgtgga	acgtgcagct	cgccgcgcgg	accatgaccg	cagtgccgat	12840

tctcgtcgtg ttcgtgctgg cgcagcggca gttcatcgag ggactcgcgc acaccggggt 12900  
taaggggtga gtcccgcgat cgacccggag atcgccgcgc tggtagacct cgccagacca 12960  
ccggtgacgc cagcgggaat cgacgcggtc cgggcgggag gacgcgtcgt caccgatgcc 13020  
gagctgacgc gcggtggcac cgtgacgttc gccgatgccg atgccgatgg cgtgcccgtg 13080  
cttgtgctgc gccagcggg tgtgccacgt cttccggtgc tgcacctgca cggcggcggg 13140  
atggtcgcgc gcacccggcg caccgacctg cacgtgctgg ccgagtgggt gtcggagctg 13200  
ggtgtcgtgc tgggtgcgcc ggagtaccgg ctgcgacccg agcaccgcga cccggcgccg 13260  
tcgcaggact gcttccgggt gctggagtgg atgtcccga acggcttcgg cccgcctgtc 13320  
gttgccggta catcggcggg cggcgggctg gcggctgcgc tgacgtgat ggcccgcgac 13380  
ctcggcggtc cgccgatctt gccgtcggcg cgcaactgct cgaacgcgc cagcgtctcc 13440  
gccaacggcg tgctaccgcg ccagtgcagc aggtagacct ccaggtgatc ggtgcccaac 13500  
cgccgcagac tgcgttcgca cgctcgcag gcgccacgcc gaccgcggt gtgcggatac 13560  
accttgcgtga ccaggaaaac ctgcgcgcgc cgcccggcga tcgcggcacc gacgacctcc 13620  
tcggcgcccc cgctgccgta catctcggcg gtgtcgatca gccccaggcc gaggtccagg 13680  
ccgcgcgcga gcgcggcgac ctccgcgcgc cgctgcccac gccgctcccc catgccccag 13740  
gtgccctgcc ccagcaccgg caattcctgc ccgccacgca gttgaaggcc tcgcatgcac 13800  
actccgtttc agtcagcgt gctgcggcag caccggccacc agctcgaagc caccgtcctc-13860  
ggacggcccc gcggccaacc gccacccggc catctgcacc cgctccgaca ggccggccag 13920  
cccggaccgc cgctgccga cctcgatggc cgggtcccgc cggtcggggg cgccgtttgt 13980  
gacggcgacg tgcagcgagc cctcgcggaa caccagccgc accttgatgt cagcaccggg 14040  
tgctgtcttg gaagcgttcg tcaacgcctc ctggacgacc cggtaagccg cgcgtcaac 14100  
cgcaggcgac acccgtgcg ggtcgcggat ctccgaggaa agctccacat ccacaccggc 14160  
ggcgcgcgcc cgcacgacca gctccgggat cgtgctcaac ccggcccgt gctgggggc 14220  
gccgagcccc agcgtcgcgc gcatctcgtc cagcgcctgc ttcgccaacc cgcgcagccg 14280  
cccggcggtc tccctgcac cgggatactc ggtggtggc gccaacgcc cgactcgac 14340  
cgcgattaag gtcacgtgat gtcccaccgc gtcgtggatc tcccgcgcga tccgggcccg 14400  
ctcctccgcg cgcgcgctct cggccttcgc ctgcaactcc gcctcgggtg ccgcctccag 14460  
cttccgcaga ctgtccgcca gctgctcccg caacgccacc agcgccccca atgcctcgg 14520  
cgcgcctgcc gccagcagca cgaacgccac ggtcaggatg atcgaccga ccccagcga 14580  
ctgcatgatc atcaccggtg ccacggcgac cgccgcggtc agcaacaccc agccgaccaa 14640  
catctggatc cgcggctgcg cgcgccccaa cgcgtacaac gccacgaccg tcggcgccca 14700  
acccaaacca ccggccaacg ctggcatgca cagcagcacc gcgatccgc gccaccgccg 14760  
caacggcagc agcgcgaag ccaccaaacc ggacaccgtc gcatagccga acgggtcggg 14820  
cgcggccagc acagtgccg ccggaatccc gacgaccagc gcctcaacgc cgagccgaga 14880  
ccagcgacc gaactaccg gaagccgcgc gccacgcgt gcgcgatcaa cgcggcctgc 14940  
acccggttct gcgcgcccag cttgctcaac accgtcgaca cgtagctctt gaccgtcgcc 15000  
tcggtcagcc cgagccggac cccgatatcg gcgttggaac gaccatccgc caacaactgg 15060  
agcactgcc gctcccgcgc ggacaacccc tgcaccagct gcgtctcgtg cgcgttgctg 15120  
gacagactgc ggaacctggg caacaaccgc gcagtgatcc gcggatccag caccgcccc 15180  
ccggcgcca gatcgtgcac cgcccgaacc agcaccgcgg gctccgcgtc cttgagcagg 15240  
aaccgtgcg ccccgaaacc caaggcatcg gccacgtaat cgtccaggtc gaacgtcgtc 15300  
agcaccgcgc gcaccggcg attcgccaac gtcgccaaat cccgcaacgc cgtcagcccg 15360  
tccctgcgg gcactgcac gtcgatgagg gcgacgtcga cgcggtacgc gcgcaccgcg 15420  
gccagcaact catccccgtt gccgcctcgc gcgaccaccc ggatcgagcc gtcgccctcc 15480  
aacaacaccc gcaggccacg ccgcagcacc gcctcgtcat cggcaagcac aacgcggatc 15540  
ggcccccgcg cgtcgtcagc ccgcacctct cggtcacctc cctgttccga gggtcgccag 15600  
cagactagcg gggccgcgaa agcgcagcgc acacgccacc cgagcgact gtgtcagcg 15660  
cgccgcctc tggccttcgc ctcccgcacc cgcaccgaga tccgcaccgg gctaccacg 15720

aacccgaacg	tctcccggaa	cctgcgctcc	aggaaccggc	ggtaccgggc	ctccaagaac	15780
ccggtggtga	acagcacgat	cgtcggcggc	cgcgactgcy	cctgggtggc	gaacaggatc	15840
ttcggctgct	tgccgccccg	caccggcggc	gggttgggcg	cgaccagctc	ggaaagccaa	15900
ccgttcaacc	ggccggtgga	gatccgggtg	tcccacgaat	ccagcgccgt	gcgcagcgtc	15960
ggggccagtt	tcgccaccgc	gcgaccggtc	ttcgccgaca	cgttgaccgc	ctcggccccac	16020
cgcacccgca	ccagctcgcg	gtcgatctcc	ttctccaact	ggtggcgggc	gtcctcgctc	16080
accagggtccc	acttggtgta	ggcgatcacc	aatgcccgc	cggcctcgac	gaccatagtg	16140
atcacccgca	ggtcctgctc	gctcaacggc	tcggagccgt	cgatgagcac	gatcaccacc	16200
tcggccgcct	cgatcgcggc	cttggtgcgc	agcgagcggt	agtactcggt	gccgcttgcg	16260
gtcttcacgc	gcttgcgcg	cccggcgggtg	tcgacgaacc	gccacacctg	gccgtccagc	16320
tccaccagcg	agtcaccgg	gtccaccgtg	gtgcccgcga	cgtcgtgcac	caccgaccgc	16380
tcttcccccg	tcagcttggt	cagcaggctc	gacttgccca	cgttcggtt	gccgaccaat	16440
gccacccggc	gcggccccacc	ggtcgcgccc	aagatctcgc	gcggcgcttc	cgggaacacc	16500
tccaggacgg	cgtccagcag	gtcaccggaa	ccgcgcccgt	gcaacgcgct	gaccggcatc	16560
ggctcgccca	acccagcgga	ccacagcgaa	tgcacgtcgg	agacgcctcg	ctggtcgctc	16620
accttggttg	cggccagcag	caccggacgc	ttggaccgcc	gcagcaccct	ggccacggct	16680
tcctcggtct	cggtggcacc	gacccgggcg	tccaccacga	gcagcaccgc	atcggcggtg	16740
tgcattggca	gttcgcctg	cgcgccacc	gacgcctgca	gtcccttcgc	gtccggctcc	16800
cagccaccgg	tgtcgaccac	cgtgaagcgg	cgccggttcc	acaacgcgtc	gtaggccacc	16860
cgatcccggg	tgacaccagg	ggtgtcctgc	acgacggctt	cccgcgccc	cagcagccgg	16920
ttcaccagcg	tcgacttgcc	cacgttgggg	cggccgacca	cggccagcac	cggctgcggc	16980
ttcgtcggct	cctcgccctc	gaccgcttcg	acctcgctga	actcagccca	ctcagcctcg	17040
tcggaccacg	tgccgtccag	gccttccacc	gactcgctcg	tcacgctcgt	ttccctctcg	17100
aatagtcttg	gctctgggtc	gcgaccagcc	gatccagctc	ggcgatcagc	tcgaccagtt	17160
cgggtgcgcac	cgcctcggtc	gccacgacca	gcccggcgcg	gcctttgccc	gccggcagca	17220
cgatcggtc	cccgaacaac	acgtcgatcc	tcggtagcag	cctacggccc	gggcatccg	17280
ggcggcgcgt	gccccggcag	gccaccggca	ggatcagcgc	ctccgacgtc	cgcgccaacc	17340
acgccgcacc	gtgctgcgca	ttggtcacat	cgccatcacc	ccgggtgccc	tccgggaaca	17400
ccccgacgag	accaccggcc	cgcagcacc	gcaccgcgc	cagcagcggg	gcgcggtccg	17460
gcgcaccgcg	gcgcaccccg	atctgcccga	tgcggcgcag	gaaccagccc	agcgcaccgc	17520
ggaacatctc	ctgcttgatc	aggaacaccg	cgtggcgcg	gatcattccg	aacagcagcg	17580
gcccgtccat	catcgaactg	tggttgggca	ccagtaccac	cgggcccgtg	cgcgggatcc	17640
gctccgcgc	gtggatccgc	accgggtacg	gcagctgcac	gatccggcgc	gaaatccact	17700
ggccgcaccg	gtgcattcca	ccggacgcgc	cctcgggcag	ggtctcgctc	ctcatggcgt	17760
cgtccgctcc	gccgcgacca	gcaggccacg	gcactccacg	tgcttgcgca	actgcgtgag	17820
cacctccacc	acaccaggt	ccgtgggtgtc	gagctccacc	gcgtcgctcg	ccttgcgcg	17880
cggcgccacc	ttccgccccg	agtccagcgc	atcgcggcgc	tgacagtcgc	cgtgggtccg	17940
gtccagggtca	ccggcccgtc	cctcgagagc	gtcctgcgc	gtgcggcgct	gcgcgcgggc	18000
gtgcgcggac	gccgtcaggt	agaccttcaa	accgcaccc	ggggccacca	cgtgcccgat	18060
gtcgcgtccc	tcgacgacca	tgccgcctgg	cgacaccagc	gcctcaccga	tcagcttccg	18120
ctggtggggc	accaaactgt	cccgccacct	ggcgaccgc	gacacgcgc	acaccgcgc	18180
ggtcacctcc	ggcccgcgga	tctcccggcc	cagtcgcgag	ccgtccagga	agatctcagg	18240
gccctccgga	tcggtgcccc	ccgtcagccc	cgctgcgcgc	accacgcctg	ccactgcggc	18300
cggatccgac	gggtcgacct	cggcgcgcg	caccgccagc	gtcacggcgc	ggtacatcgc	18360
tccggtgtcg	aggtaggtag	cccccaatgc	gaacgccaac	ttgcgggaca	ccgtcgactt	18420
gccggtgccc	gaaggggcgt	cgagcgccac	cacaccaagg	agcctggcgt	gtgccacgtg	18480
catgtccttc	ccgagccgga	accgatcggc	ccgacaccgc	gccggaccaa	ccgtccattc	18540
tgcctggtcc	gctcctgcgc	atcccaatcg	catccagatc	caatccggac	tgccgggacaa	18600

atacgaaagt tctacgaata aggtcttggg tgtggatgta actgtgaccc ccttgccagg 18660  
 gctcggcacg cagcaggact tcaccacccg atccggccac cgcacgcg cgatcaccta 18720  
 ccgggacggc cgcttcgagc tgatcgtctc cgaccacgaa gacccagaca aggtggccgc 18780  
 ctccggtcgcc ttgaccacct ccgagaccag cacgctggcc aacctgctgg gcgccccgca 18840  
 gctggtcgcc cagctcaacg agcagcacag cgaggctcgcc gggatcacca cctggcagct 18900  
 gtccggtcgcc ccggggtccc cctacgcgcg ccgaaccctc ggcgacacgg agatgcgcac 18960  
 ccgcacctcc gtgtcgatcg tcgcggtcgt ccgggacggc accgtgcacc cctcaccccg 19020  
 gccggacttc cagtctctcg ccggcgacct ggtggtggtc gtgggcaccg cgaagggggt 19080  
 acgcgcagcc agcgaaatcc tggaaaaggg ctgatccggc ccgtcccga cactttcccg 19140  
 gccacctgac cgggacacgc gtcaccaacg gcggtccagc gcgtgccgat cccctaacaa 19200  
 ctccatacgc agggcacacg cgagaggaaa cacgttgcaa gacacggcga tctccctaata 19260  
 agaactgggt gcggttttct tcggtctggg cactctcggg agactcgcat ggaaaatcgg 19320  
 ggtgtcaccc atcccgctct acctgatcg cgggctggcc ttcggcaccg gtggcctggt 19380  
 cccctgcac ggcatcgaac cgttcaccca cctcgctcc gagatcgggc tggctctgct 19440  
 gttgttgctg ttaggcctgg aatactccgc ggggtgaactc gtcaccggct tacgcgcctc 19500  
 ctgggtggcc ggtgtcatag acatcgctgt gaacgcggta cccgggtgcgc tgggtggcct 19560  
 gctgctgggt tgggggtccgc tcggcgcggt cacaatggcc ggcgtcacct acatctcttc-19620  
 ctccggcatc atcgogaagg tgctcggaga cctgggtcgg ctcggaacc gggaaacgcc 19680  
 ggtgatcctg tcgatcctgg tcttcgaaga cctggcgatg gccgtctacc tgccgatcct 19740  
 gaccgcctg ctggccgggg tcagcttctt cggcgactg accgcggtcg ggggtgcgt 19800  
 ggtcgtgatc acctggtgc tgggtggtgc gctgcgcttc ggcaagtacg tgtccgcgt 19860  
 ggtcgacagc ccgcaccg aggtcttct gctgcgctg ctggggggcg cgtgctcgt 19920  
 cgcggggatc gcctcggagc tgcaggtctc ggcggcggtc ggctcgttcc tgcgggcat 19980  
 cgcgatctcc gggtcgaccg cggcgaaacgc gacccgcagt ctggagccgc tgcgcgacct 20040  
 gttcgcgcg ctgttcttcg tcgtgttcgg cctcaacacc gatccgagcc agatcccgc 20100  
 ggtgctgccg gtggcgtgc tgcgtggcgt ggcgacctgc gcgacgaaag tgctcacccg 20160  
 ctggttcgcc gcccgatgc aggggtgctg ccggatgggt cggctgcggg ccggtgccgc 20220  
 gctggtggcg cgcggcgaat tctccatcgt gatcgccggg ctggccgtcg cctccggcg 20280  
 ggtgccgggc gagctcgccg cctcgccac cgcctacgtg ctgctgatgg cgatcgtcg 20340  
 cccggctcg gcgcgcgtcg tcgagccggt ggcccgatg ttcctccca gccaggcgaa 20400  
 gaaggcctga ccggaaaagc ctccggctgt cccgagcgga cgggcccggg gctttctcaa 20460  
 cgtgcggat aacgcggcg ggtctggcaa acgcaacacc ggtgcgcgaa cagcctcaga 20520  
 ggccactgc gcggtagagc gagccgatct cctggcggtt caacggccgg atcgcgccc 20580  
 ggcgtcgtt ggtcagccgc acatcgccga tccgcgtcc caccaggcg tgcaccgggt 20640  
 agccgacgtg cttgagcagc cgcgcacga tgcggttgc gccctcgtgc aacacgatct 20700  
 ccaccagcga ccggttctgg ttcattgctg ccagcttgaa ccggtccacc ttcaccagac 20760  
 cgtcctccag ctccacgcc tcgcgcagct gctgcccgag gtccttcggg atcgaccca 20820  
 gcacctccgc caggtacgtc ttgcgcacct tgtacgacg gtgcacagc cgggttcgca 20880  
 gctcgccgtc gttggtgate agcagcagg cctcggtgtc ctggtccagc cgcgccacgt 20940  
 ggaacagctt gccctcccgt tccgcagggt aatcacccat gcacggacgc ccttgcgt 21000  
 ccgacatggt gcacaggatg ccgcgcgggt tgttcagcag caggtgcgtc agatcgtcgt 21060  
 tgaccatcac ccggttgccg tccacgtgga ccaccgcgt cgtcggtcg accggcgcc 21120  
 ccagctcggg gacgacctc ccgtcgacct cgatccgcc ctcgacgat atctcctcg 21180  
 ccgcgcgcc cgaggccacc ccggcctggg acaacacttt ctgcagccgc acgccttcg 21240  
 aatcggcagt gctgctgtac ggggaacggt gctcagacgt catcaatga atccacttcg 21300  
 ggcaacaagg ggcgagcgg cggcaactcc ttcaacgacg acagccccag ccgtccagg 21360  
 aacagctcg tcgtgcagta caggatgcc cccgtctcc ggtcggttc gccctcttcg 21420  
 ataaggccgc gcccaccag ggtacggatg acgcgctcga cgttcacacc ccgtactgcc 21480

gcgacccgcg aacgcgtcac cggctgcgcg tacgcgatca ccgcgagggg ctccagcgcg 21540  
 gcacgcgtca gcttcgaacg ctgcccgccc aacaggtacc gctccacgta cggggcggtac 21600  
 ttctcccgcg tgtagaaccg ccagccgctcg ccgacccgcc gcaggtcaac gccccgttcc 21660  
 gcctcggcgt agccgtccga gagccgcgcg agcgcggacc ggatccgggc caccggctgc 21720  
 tccaggggtg cggccagcag ctctcgcgcg gccggcacgt cgaccacgag cagcaacgcc 21780  
 tccagcgccg catccagcgc cgaatcggag gtcaggtccg gcgcactcgg cgcctccacc 21840  
 gcctcgggaat ccccagcggc cgccgacggc tccaacgggt cctgcggctc agcgggcggc 21900  
 ggggtcggcg actcctccgc gacctcgcgc ggggcctccg gctccagctc cgtcaccctg 21960  
 actcctcttc ttcattggctg gcgcgctgcg cctcggcttc ggctccgcgc tgcgccacgc 22020  
 tgccgcccac ccagctcacg atcagctcac ccagcggggg ctctctgctcg aacagcagga 22080  
 ccttctcccg gtagagctcg agcagcgcca ggaaccgcgc caccacctcg acgggtgtgc 22140  
 cgcagtcgga caccagctcc gagaacgtcg cgtgcctctt ctccgccaac cgcacccgca 22200  
 gcagcgccgc gtgctcgcgc accgaaacgc cgtgctggtg gatgtggtcc agcgaacagg 22260  
 tcggcggtgg cttcggccgg aacaccgcgc ccgcgatctc ggcgaaaccg tgcggcgga 22320  
 ccccgatcac cacttcgggc agcaggttcg cgaaccgcgc ctcgaccgac accgaacgcg 22380  
 gatagcgccg cagcgccccg gcctccaact cgccgaacag cgcggccacc tgettgtacg 22440  
 cgcggtactg cagcagccgc gcgaagagca ggtcgcgggc ctccagcagc gccaggtccg 22500  
 cctcgtcttc gacctcgcgc gccggcagca gccgggcccgc cttcaggtct aacagcgtcg 22560  
 cggcgaccac caggaaactcg gtggtttcgt tcaggtccca ccgctcgcgc agcgccttgg 22620  
 tgtaggcgat gaactcgtcg gtgaccttgt gcagcgctac ttcggtaacg tccagctggt 22680  
 gctgcgagat cagctgcagc agcaggtcga aggggcccctc gaagtgttcc agccgcaccg 22740  
 tgaaccggcc gcctcggcg ggatcctcgc tcgactgcgc cgcgggtacc gcggggtctg 22800  
 tcaactctcc ggctcctgca ggcgcgcgc cagcaccgag tccgcgcgt tcttgtcgag 22860  
 gtcggcgagc accacggcaa ccgcttcgcg caccacgcgg cctcgggtcca cggcgatgcc 22920  
 gtggtcggcc cgcaatgcca gccgcgcctg ctccaagccc agcaactcct cgtcggacac 22980  
 gtacacgggt atcttcgaat cgtgcttgcg ccgtccagag ccgccccggc ccgcgcggg 23040  
 tgccgcctgc tccgggtccg gcggattcgc gtgatccgtc cgaggaacag tagtacgga 23100  
 caattcggcc gctccgggca gggcgacccg acgcgtcctc gggcgatcac ctgcgcgcgc 23160  
 agtgcgcggt acgcctgcgc cccggccgag cgggggcgcc agcgggtgat cggctcgcgc 23220  
 gccaccgtgg tctcgggaaa ccgcaccgtg cggttgatca ccgcacgaa gacgatgtcg 23280  
 ccgaacgcct cgacgacgcg cgccatcacc tcgcgcgagt gcagcgtgcg cgggtcgaac 23340  
 atggtcgcga ggatgccgct gatctccagc ctccgggttc gccgctcccg cacttctcgc 23400  
 atggtgtcga tcagcagcgc caccgcgcgc aggtgaaga actcgcactc cagcgggatt 23460  
 atcacggcgt ccgcggcggc caacgcgttc accgtcagca gcccagcga cggctggcag 23520  
 tccaccagca cgtagtctga ctccgccagc gccggctgca gcacccgggt cagcgtctgc 23580  
 tcccggccca cctccgcgac cagctgcacc tcggccgcgg acaggtcgat gttgctgggc 23640  
 agcaggtcca tgccctcgac ggtggtctgc cgcaccacgt catggatggc gaccggccgt 23700  
 tccatgatca cgtttagat ggtctggtcg agctggtgcg gctggacccc gaggccgacc 23760  
 gagagcgcg cctgcggatc gaaatccacc agcagcacc ggccgccgta ttcgccagc 23820  
 gctgcgcga ggttgatcgt cgacgtggtc tccccagc gcgcctctct gttgcacatc 23880  
 gcgagcacc atgcgggacc gtgcgggtcc agcagcggcg gctcggggat gtgcggacgt 23940  
 ggccggccag tgggaccgag gccgcggggg ttccggccca ggccgaacct ctctcgcgcg 24000  
 cccgcctct cccgggcgat gctcaggtcg accgcgcccc gggaccgccc ttcggcggcg 24060  
 ggctcgggca gcgacatggc gatctgactc cttttggtgc cgggtggcgat caaacgcagc 24120  
 ctaagtggca ttcgttacca gcggcaacgc gcctcgcgt ggtgtcggaa gttccctgtg 24180  
 atctgcgtgt gatcaacgtc taccgcgtac accacccct cgcattcgca gcgtcaccaa 24240  
 gtcactccgt agcgtgtgcg cgcggatgcg tcgtcgcgta gacctccgc agcgtgttca 24300  
 ccgtgatcaa ggtgtagacc tgggtcgtcg tgaccgagc gtggcccgag agttcctgca 24360



ccacgcggac atccgcgccc cctccacaa ggtgctggc gaacgaatgc cgcagcacgt 24420  
 gcggcgaaac cgagccgttg atcccgccgc gctgcgcggc ggtcttcaac gcgttccacg 24480  
 agctctgccg ggaaagccgg gtaccgcgcg aattcaggaa caccgcggca ctaccccgcc 24540  
 cgcgcgcggc caggccgggc ctggcccgtg ccaggtagcg gtccagcgcc gccagcgccg 24600  
 gtcggccgat gggcaccagc cgctggcggc cgcccttccc ttccagcagc acggtccggt 24660  
 tgggtgccgtc gatgtcgtcg aggtccaacc cgaccgcttc ggagatccgc gctcccgctg 24720  
 aatacagcac ttccagcagc gcccggtccc gcaggccccg ggcatcctcg ccccgggcgt 24780  
 gatcgagcag cttcccgacg tcgccgatgg aaagcgctt cggcagccgc ttcgggcgggc 24840  
 tcggcgggcg gacatcgcg gctacgtcga ccggcagcag cccctcgggc tgcgcgaagc 24900  
 ggtgcaggcc gcggcgggcc accaaagctc gggccgacga cgacggcgcc agcgcgggcc 24960  
 gctgcccggg gccttcccgg agctccgcca agaaccggga cagggtgctc gaacgcacct 25020  
 cgccgagccc agaaatccct gccgcgacca ggaattcggc ataccggcg aggtccctgg 25080  
 cgtaggagtc gagagtgtcg cggcggttac cgcgctcgac ggcaaggtgg tcgaggtacc 25140  
 cgggtgatcg cccgcgcagg tccggcggca gctcctcgaa acccctcacc gggcgggcga 25200  
 gcgctggggc cgggtccggc attcggcac cgccggggcg ggctccacc ggccctcgcg 25260  
 gaccgcgtgg gcggcgagca gccccgccac cccggggcgg ttgacgatct ccccgggccag 25320  
 cgccatccgc accgcctcg cgagcgggaa gcggcggtac -25360

<210> 4

<211> 45624

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 4

tcagccgacc ggcttccgcg ccgtcagaat cgcatagcc atgtcgtggg catatttctc 60  
 ataatcgag acccgggcgg cccagccggc gacagccggc ccgtacctgt ccgcgatccc 120  
 gtgctggtgc gcagcgagct cttcggcgaa ctgcggcatg aagtaccggg tctcgcagca 180  
 cagtcctcc caatcgagga tctcgaaccc cgctgcacgc agcgattcca gaagtgtctc 240  
 agccaggcag atccgaaggc cggtcggcca cctgtccccg gacaccggca tcccgccgcc 300  
 cgcttctcgt ttgacgacct cggtagcgcc gaggatgcca ccgggtttga gtactcgaag 360  
 gatttcccg atggcacggt ccggttcgga catctccaac agcgactgca tggcccaggc 420  
 ggcgtcgaaa gcattgtccg ggtacggcag ggacatggca tcgacgcacg agaagtccac 480  
 ccggtggctt agtcccgctt cgcgtgcgca atcagcggcg atggccactt gcacctggct 540  
 gacggtgatg ccggtgatct ggatcgcggt gtcgcgcgcg acgcgcagcg ctggttgctc 600  
 ggtaccgcac cccacatcga gcagtcgaac gccgccatcg agcacggctc gttcggcgac 660  
 aaggtcggtg agccggtcgg cggcctgctg ccaggaagcc cgcccgctgt tctcccagta 720  
 gccgtggtgg atggcgagg ggccgcccgc gaccgagttc agcaacggcg tgaccaggtc 780  
 atacatctgc ccaacctgct gcgatgttg tgcgccacct ggcaacaccg gtatgcctga 840  
 tctctgcaac gttcaccttc tcgtaatttt ctggaaccga gcaggaaaca cacagagaga 900  
 aaagacggca aactagccgt cagttaatto gcggttaccg ccgcattcgg cggtaccgac 960  
 tataactatc cggccaggag gcaaacagcc agcgcaaatc ttcccatga atcccgcgag 1020  
 aaccaacggt gatcacgta caggtcgggc ccaccccgat cgagcggagc tcgcgcattt 1080  
 cactcgttcg cgtgaaacca ggtgcggcga ccaaagatc ttcgatggtc gcgtgacatt 1140  
 cctgccgtga cacactattg ttgccgcgg aagtcggttg gacgtcaacg gcgatcgaca 1200  
 tacgcagcgc attttccaaa tctggccaa tgggagtagg ggctgctgcg ctaatcgggg 1260  
 gaaagagact ggggtccgaa gtatgcgcgt actcgtcgtt cccttgcct atccgacgca 1320  
 tctcatggca atggtgccgc tgtgctgggc gctgcaagca tccgggcacg aggtcctgat 1380



cgccgcacca ccagagctgc aggcgaccgc gcatggtgca ggtctcacca cggccgggat 1440  
ccgcgggaac gacaggaccg gcgatacggg tggaaaccacg cagctgcgct ttcccaatcc 1500  
ggcgttcggg cagcgcgaca ccgaggcagg ccggcaactg tgggagcaga ccgcgtccaa 1560  
tgtcgcgcaa agctcgtctg atcagctccc cgaatacctt cgactggccg aggcctggcg 1620  
accgtcagtg ctgttggtcg acgtctgcgc gctgatcggc cgggtgctcg gcggattgct 1680  
cgacctgccg gtcgtgctgc accgctgggg agtcgacccc accgcaggcc ccttcagcga 1740  
tcgagcccaac gagttgcttg acccgggtgtg ccggcaccac ggactgaccg gcctgcccac 1800  
tcccagctc atcctcgatc cctgtccgcc gagcctgcaa gcaagcgacg cgccgcaagg 1860  
cgcaccggtc cagtacgtgc cgtacaacgg aagcggcgca ttcccggcat ggggcgcggc 1920  
gcgcacctca gcacggcggg tctgcatctg catggggccg atggtgctga acgccaccgg 1980  
gccggctccg ctgctgcgcg cagtagcggc tgccaccgag ttgcccggcg tcgaggccgt 2040  
gatcgccgtt ccccccagac accgggcact tctcaccgac ctaccgcaga acgcccgat 2100  
cgccgaatcg gtcccgtca acctgttctt gcgtacctgc gagctggtca tctgcgcggg 2160  
cggtcggga accgcattca ccgcgacccg actcggcatc ccgcaactcg tcttcccca 2220  
gtacttcgac cagttcgact acgcgcgcaa cctcgcgcgt gccggggcgg gcatctgctt 2280  
gccggatgag caggcccagt ccgaccacga acagttcacc gactcaatcg caacgggtgct 2340  
cggcgacacc ggcttcgctg ctgcggcaat caaactcagc gacgagatca cggccatgcc-2400  
ccatcccgcg gcgctggtgc ggacgctgga gaacactgcg gccatccgtg cctgatgaac 2460  
tagtcacccg agaacagctg gatccggaga accgatgcc tcccagaacg cgctgtacct 2520  
ggacctgctc aagaaggtac tcaccaacac gatttacagt gatcgggcgc atccgaacgc 2580  
ctggcaggac aacaccgact acaggcaggc cgctcgggcc aaaggcacgg actggccaac 2640  
tgtcgcgcac acgatgatcg gtctggagcg gctggacaac ctccagcact gcgtggaagc 2700  
cgtgctcgca gacggtgttc ccggggattt cgccgagacc ggtgtctggc ggggcgcgcg 2760  
atgcatcttc atgcgcgcgg ttctccaggc attcggagat accggacgta ccgtctgggt 2820  
agtggattcc ttccaggga tgcggaaag ctctgcgcaa gaccaccaag cggaccaggc 2880  
tatggcgctg cagcagatca acgacgtgct tggcgatatc cttgagaccg tccggcagaa 2940  
cttcgcccgc tacgggctgc tcgacgaaca ggtcagggtt cttcccggct ggttccggga 3000  
caccttgccc accgccccca tccaggaaact cgccgtgcta cgactcgacg gcgacctcta 3060  
cgaatccaca atggactcat tgcggaacct gtaccogaag ctctcgccgg gcggattcgt 3120  
catcatcgac gactattttc tgccgtcctg ccaggacgcg gtgaaggggt tccgcgcgga 3180  
actcgggatc acggaaccca tccacgacat cgacggcacg ggcgcctact ggcgcgcgag 3240  
ctggtgaacg gtcagccgt cctcgacgcc gagcgcttgc cggggcacga accccggcgc 3300  
ggcaggctca gcgttgagcc ctttctccac gaataccagg ttgtggtaga agtgcagggc 3360  
cgccacgttc cgttccgtgt aggagggtc ggtcccgcac cgcgattcgc gctcctgata 3420  
atgcaggccg tcgatcagtt ctttgagcat gtcgatcgag gtgcgctggg ccgcgggttc 3480  
cccacgcga ccgcgtagc cgggccagta cgccgtctgg agatcctcga tgacgtacaa 3540  
accacctggg cggacgtgcg gaaacaggga ttggaaggat ttcttgacgt ggtcgttgac 3600  
atggctgccg tcgtcgatga caatgtcgaa cggcccgatc tttcccgcga tgtcttcag 3660  
gaattccgca tcgtctggt cacctcgag ctttcgact cggtgccctt cgttcccggc 3720  
tttctcgaat atgtccagcc cgtaaaccg acctcgccgg aagtaccgt gccacatgcg 3780  
cagcgaagca ccaccgagtt cgggtgctg ataaccacc attcctatctt ccaacactcg 3840  
caccggggca tccctggaatc gggagaagt gtgctcgtag tgttcgggt accagtgcag 3900  
gtccgcccac ttgtcggatc cgtagcggac tgccagcgca ccgaggtcg cgcgccgtgt 3960  
ggcggctgcg gccagcacgg cgtgcaccgc ggtggacacc ctgttccgga acgccagcag 4020  
gcgctgcgcg ccggcgaggc cctggctcgg cgagaactgc gtcatgctgt ccgaccagcg 4080  
gacttcacgg ctgttgtgcc gcctgccgtc aaccgggccc aagagccctt ccagcagatc 4140  
caccgcgtcg aaccggagaa cagcagctgc ttcggcgacc gccgcgagcc gcaggcctgc 4200  
gtgatcgagc tcaacggttc tccggaccag ctgagcgcca gaagtgatct ccaggccgat 4260

tcgcaccgac tcgtccaacg acagcggatc gcagcgtacg acgagttcgt cgacgatggc 4320  
 gtcggccacc gcgtccagtc cggccacctg cactgcttcc tggagcttcg ccgggccccg 4380  
 acccgccgag agcagcaaac gctccaccac cgaccagggg gcaactgcga tctcactcat 4440  
 ggaaagtcat caccttttcg gttccagcgc atacggcacc atgcgcaccg cgatatcgat 4500  
 ccacaaagta gagccggcgg aggcgaccgg gtttcacgta tgagctgaaa cacacagata 4560  
 atcaccggga cgcgtagatg atctcggctg cgggcgaaca aagtggacca gtcagaaaag 4620  
 gaggggaggc gcccgaaatt catgacccgg caccatgaa tcgtcgaacc ccaggaacag 4680  
 agatcaccgt cgagcccgac gatcctcggt atccggacct cgtcgtcggg cacaaccccc 4740  
 gtttcaccgg aaaacccgaa cgcattccaca tcgccagctc cgccgaagac gtcgtgcacg 4800  
 ccgtcgcgga cgcgtgcgc accggcaggc gggtaggggt ccgcagcggc gggcactgct 4860  
 tcgagaatct cgttgccgac ccggcgatcc gagtgtcgt cgacctctcc gagctcaacc 4920  
 gcgtgtacta cgacagcacg cgcggggcat tcgcgatcga ggccggcgcc gccctcgggc 4980  
 aggtgtaccg aaccctgttc aagaactggg gcgtgacgat cccgaccggc gcatgtcccg 5040  
 ggggtggcgc aggcgggcac atcctcggcg ggggatacgg cccgctgtcg cgcgattcgg 5100  
 gttcggtcgt cgactacctt caaggcgtcg aggtcgtcgt ggtcgaccag gccggtgaag 5160  
 tgcacatcgt cgaggccgac cggaaactcca cgggcgcggg tcacgacttg tgggtggcgc 5220  
 acaccggtgg cgggtggcggc aacttcggga tcgtcaccag gttttggctc cgaacgcggg- 5280  
 acgtggtcag caccgacgcc gcagagctcc tgccacggcc gcccgcgaca gtgtgtctcc 5340  
 gatcgttcca ctggccgtgg cacgaactga cagagcagtc attcgccgtg ctctacaga 5400  
 acttcggcaa ttggtacgag cagcacagcg cgcctgaatc cacgcaactc gggttgttca 5460  
 gcacgctcgt ctgcgcacac cggcaagctg gctacgtcac gctgaacgtt cacctggacg 5520  
 gcacggatcc gaacgcggaa cgcaccctgg ccgaacacct gtcggcgatc aacgcccgag 5580  
 tcggcgtgac tccagccgaa gggctgcggg aaacctgccc gtggttgcca tcgaccagag 5640  
 tggccggggc gatcgccgaa ggcggcgaac cgggcatgca acggaccaag gtcaaagccg 5700  
 cctacttgcg caccgggctg tccgaagctc aactagccac ggtttaccgg cggctgaccg 5760  
 tctacggata cgacaacctt gcggcgggcg tgttgctgct cggttaccgg ggtatggcga 5820  
 atgccgtggc tccgtcggcc accgcactcg ctacgcgcga ctcggttctc aaagcgtgt 5880  
 tcgtcacgaa ctggtcggag cccgccgagg acgagcggca tctgacctgg attcgcggtt 5940  
 tctaccgcca gatgtacgcc gaaaccggcg gaggttccgg gccaggtacc cgtgtcgacg 6000  
 gctcctacat caactaccgg gacaccgact tggccgatcc attgtggaac acctccgggt 6060  
 ttgcctggca cgacctgtac taaaaagaca actaccgcgg gctgcagcgg gccaaagcgc 6120  
 ggtgggatcc gcagaacatc ttccagcacg gcctgtcgat caaaccgccg gcacggcttt 6180  
 cacccggtca gccatgagga gctcatcacg atgtccacaa cgcacgagat cgaaaccgtg 6240  
 gaacgcacat tctcgcggc cggatccagt gcggcgagcc tggccgacct gaccaccgaa 6300  
 ctcggaactc ccaggatcgc acccgtgctg atcgacgaga tctcttccg cgcggaaccg 6360  
 gcccccgaca tcgaacggac cgaggtcgcg gtccagatca cccaccgagg cgagaccgtt 6420  
 gacttcgtcc tgacgctaca gtccgggtgag ctgatcaagg ccgagcaacg accggtcgga 6480  
 gacgtccgcg tgcggatcgg ttacgagctc accgatctca tcgccgagtt gttcggccca 6540  
 ggagctccca gggccgctcg cgcgggagc accaacttcc tccgaaccac cacatccggg 6600  
 tcgatacccg gtcgctcgga actgtccgat ggcttccagg ccatctccgc agtggctcgc 6660  
 ggctgcgggc accgacgtcc cgacctcaac ttgctcgcct cccactaccg cacggacaag 6720  
 tggggcgggc tgcaactgggt caccocgcta tacgagcgac acctcggcga gttcgtgat 6780  
 cgcggggtgc gcatcctgga gatcgggtgc ggtggctaca acttcgacgg tggcggcggc 6840  
 gaatccctga agatgtggaa gcgtacttcc caccgcggcc tcgtgttcgg gatggacgtt 6900  
 ttcgacaagt ctttctcga ccagcagagg ctctgcaccg tccgcgccga ccagagcaag 6960  
 cccgaggagc tggccgcggt tgacgacaag tacggaccgt tcgacatcat catcgacgat 7020  
 ggcagccaca tcaacggaca cgtgcgcaca tccctggaaa cgctgttccc ccggttgccg 7080  
 agcggtgccg tatacgtgat cgaggatctg tggacgacct atgctcccgg attcggcggg 7140

caggcgcagt gcccgccgc acccgccacc acggtcagcc tgctcaagaa cctggttgaa 7200  
ggcgttcagc acgaggagca gccgcattgc ggctcgtagc agccgagcta cctggaacgc 7260  
aatttggtcg gcctccacac ctaccacaac atcgcggttc tggagaaagg cgtcaacgcc 7320  
gaaggcggcg ttcttgcttg ggtgccaagg agtctggacg acatattgca cctggccgac 7380  
gtgaacagcg cggaggacga gtgaacagca gaggggcgaa cacacaggca tttccgaccg 7440  
cggatcaggt ggagtccttc ttgatgcgt tggcgacgg gcgtcccctg caccacggtt 7500  
actggcggg cggttatcg gaggatgcc gtgccacacc gtggtcggat gctgccgacc 7560  
aactgaccga cctgttcctc gacaaggccg cgctccgtcc cggagcgcac ctgttcgacc 7620  
tgggctgcgg caatgggcag cccgtagtcc gtgcggcatg cggcagcggc gttcgagtca 7680  
ccggaatcac cgtgaacgcc cagcatctcg ccgcccacc caggctcgcc aacgagaccg 7740  
gactggccgg cagtccttag ttgatctag tcgacggcg ccagctgccc taccgggacg 7800  
gtttctttca ggccgcatgg gcgatgcagt ccgtcgtgca gatcgtggac caggccgccc 7860  
cgatccgca ggtccaccga atcctggaac ccggcgcccg gttcgtcctc ggagacatca 7920  
tactcgggt tcgactcccg gaagagtacg cggcggtttg gacgggcacg accgccata 7980  
ccttgaacag cttcacggcg ctggtcagcg aagccgggtt cgagattctc gaagtcaccg 8040  
acctcacggc acagaccagg tgcattggtt cctggtacgt cgacgagttg ctccggaac 8100  
tcgatgagct cggcgccgc gagcctgcgg ctgtcggcac ctaccagcaa cgctacttgg 8160  
gagacatcg gccgaagcac ggaccgggac cagcacagct gatcgcccg gttgcggaat 8220  
accggaaca tccgattac gccagaaacg aggaagcat gggtttcatg ctctgcagg 8280  
ctcgaagaa gcagtcctga tggcctccga gcacgccagc ctggtcggcg acgatctgcg 8340  
ggcaccgcg gatgatccct tctaccgacc gccgacgccg ctaccgcccg gtgtcccggg 8400  
cacgtcctc agggcccgcc ccgtctcgcc actgcgcggc acgggcgaac ccgtcgcagc 8460  
caaggcctg caaatcctt accggtccaa ctccgccctt ggcatgccga acgccgtctc 8520  
cggcaccgtt ctggtgccga acatcccgtg gccgcgcgaa gatcgcccca tcatcacttt 8580  
cgcagtggg acccagggc tcggtagcca agttgccccg tcgtacctgc ttcgaaccgg 8640  
aaccgagcc gagaccgagc tgatcgccgt gccctcgac cgcgggtggg ccgtggtcat 8700  
caccgactac gaaggcctcg gtactcctg aaccacacc tacaccgtcg gcaggcgca 8760  
gggacacgcc atgctcgatg ccgcccgccg tcgcgaacgg ctaccgggct ccggcctgac 8820  
gaccgactgc ccggtcggca tctggggcta tgcgcagggt gggcaagcgt cggccttcgc 8880  
cggcgaactg caccacact acgcacctga actgcgaatc cgcgtcgcgg ccgcaggtgc 8940  
ggtgcgcatg gatctgctg acatcatcca ccgaaatgac ggggtgttca ccgggcccgt 9000  
gctggccggc ctggtcgggc atgcgcgtgc ctaccccgat ctgccattcg acgagcttct 9060  
caccgaagcg ggtcgtaccg ccgttgatca agtgcgcgag ctcggtgcac cggagctcgt 9120  
caccgcttc ctcgcccgcg agctgagcga ctctctcgac acttcgggc ttttcgagca 9180  
acctcgatg cgcgcacgac tggccgaaag cgtcgcaggt aggaacggtg gcccggtggt 9240  
ccccacgct gtctaccaca gcacggacga cgagatcgtt ccgttcgcat tcggcgagcg 9300  
actccgggac agctaccgcg cggcggttac gccagtgcgg tggcatccgc tctccgatt 9360  
ggctcacttt cccgccgcc tggccagctc gcgagtggtc gtctcgtggt tcgacgagca 9420  
cttctccgag ccgtccgca tcagcggctc gcgagatgcc aggtgagcg atggcggtga 9480  
gcctctccag cggcttccc gctggttcc ggatgcccg tcaggcttcg aggcgaacta 9540  
ctaccgccc atgcccggcg gctatgtgga ccgcaccca cgggctgac cgatcgctc 9600  
caccagcgc gcctgctgga tcatcgattc gccgaatct ccggccaccg cagggtcgtc 9660  
cacgcctgt cctgctctga tgcgcgtgc gaaggcggtg accgccttc gaacctgatc 9720  
ttccgctggc aaggacaact cgtcgacaac gcccttcgc tcgattcgga tcacggcctg 9780  
ccactcggcg ggcgcggtga acgccggctc gacgaagatt cggccacgac tccccacag 9840  
ctcgtacgc ctgcggtagt ggtgcacgaa accgtatcc aggtgggcaa cggtgccacc 9900  
ttccgattgg agcagcacgc tggccgaaa gtcgacgcc gactcctgg cctcgtgcga 9960  
gcttgcgcc agaaccgtga gcggaccgag gagaaagtgc cgagcggcac gggcgggata 10020

gacaccgatg tccagcaacg ctccgccacc gagttcgggtg cgatagcgga tgtccgtgtc 10080  
gggaagcggc ggaattccga acacggcggg gaactccgg agctcaccga tctccccgga 10140  
ttgcagcagg tcgcggacca cgtcgtgccg gccgtgggtg aggaacagg aattctcccg 10200  
cagcagcagg ttcttctctc tggccagccc gaccaggcga gcggtgtcgg acgccgtcgt 10260  
cgtcagcggg ttctccgcaa gcacgtgttt gtctgcctca agcgccttgc cgatccactc 10320  
tgcattgatg ccaggcggca acggcacgta gacggcatcg atgtccggcc gctccaggag 10380  
ccgctggtaa cccagcaccg cctcgcattc gaatcgcgct gcgaaccggt cggttttcgc 10440  
cggatcacgg ctccgccacc ccaccacctc tgtttcggcc acgtcgcaca tcgcgggcag 10500  
catccgtcgc cagcgaagg aagcgcaccc gagcacaccg atgcgcaccg gctttcgcgt 10560  
cgagctggtc atcgccccaa cgcacacaag ctatgcaggg aggcgaccaa gctgcgcgcc 10620  
tggatgttca agaagtgggt gctccggagc agctcgccca actgtcccaa agtcatccac 10680  
cggaagtgcg tgggagggtc tgcgcgaag tcctcatgca cctcgatgat ccggtacctg 10740  
ttctgcgcct ggtagaaccg accgccttct tcggacagga tcgattcgta ccgcacgggt 10800  
tcgggatcgg cggtagcac gtcgtccaag aacggcggcc agtcgttcga cggggtgctt 10860  
tggtagtgg ccacttgca ctggaccgtg ggagcgattt ccgcagtcca cttataacca 10920  
gcctccacc gagcgtggac caaacgtgc agcactccgc cgatccgctt gaccaacagt 10980  
gcgatctcac cttgttctcg cggttcgatc atcggtcgcg tccagctggc gacctacga-11040  
ttggtcgcgg acaccgacac cgcgatcacc gcgaagtact tgccgtcctg gtgggcgcatc 11100  
tcggtatcgg tgcgatacca cttgtcgacc ctgctcagcg gaacgcgcgt tgcccgcaag 11160  
cggtagcggg ccttggtctc ctcgaaacca ccgaccgcct cggagatact cgacgaatcg 11220  
atgccgtgcg agagcgacct ggccaccgcc tgtcgggaagg gctcctccga ggcggctagt 11280  
tcgggtccgg tggcggaatc gtggaacggg atgcaagaca gcaccgtccg ggtgtccatg 11340  
ttgacgatgt tgtcctgacg aaggagagcc agcacctggc cgagcgtcaa ccagcagaag 11400  
tctggcagga ctggcacttc ctctggact tccaccacca tgttccgggt gcgcttccgg 11460  
tagaaccagg acccctgttc agactggagc acgtctacca gcacgcggcc gcggccccgc 11520  
ccaggaagt agtccacata gggcggaacg ctgccacggg gtgcctgcgt gtagttgctc 11580  
cgagttgcct gaaccgtcgg cgagagctgc aggacgttga cgttgcccg ttccatcttt 11640  
gctgacatga ggcagtgcag cagccgtcg atctccttga cgagaatgcc gaggatacct 11700  
acttcagcct ggttgatgat cggctggtgc cagctgggtg ccgcgccata gttggtctcg 11760  
acctgcaggc cttctaccgt gaagaacct cgcgcggcat gaaccagggt ctccgtgctg 11820  
gcattgaatt tccatttcga caggcggtcg aacgggatgc gattggtctc gaagctgttc 11880  
tcgcccagcc gatcggcgag ccagcagtg aactccgtgg tcggaatcat gccattgcaa 11940  
gcgtcagcg cagagtcgac gaaccgcgc gtgttggttc tgccgagcgg cgcagcagca 12000  
cttgctcag cttcgacaga actgctcatg cccaatccct cactgccgtg acgatgtgca 12060  
cgccggctcc caggggcgtc gcggtctggt ccagcggcct ggcgaggatc atcgaagcga 12120  
tccgccgta aggaccgca gcacctcgtc ctggcggttc gtcattgaaga aatgcccgcc 12180  
ggagaagaca cggacgtcgg ccacggcctc ggtgtgctct cgcagggctt ccgcctcatc 12240  
caagggtgacc ttcggctcgg catctaccac cagcacggtg atcgggcaac ccaccttcgc 12300  
tcctggcgca cagcggtatg tctcgaccgc acggtagtcg ttgctcagcg ggcgccacga 12360  
tcctcgtcgg aaatcgggtg gagcgcaaac ctctggggag atcaacaccc cattggacga 12420  
ttctggacaa gaatctgact tcgggactcg cctggcatcg tcccggtcac ggcggaatcc 12480  
gagtcagcgc aacagcgatg aacacctgcc aaaccccgcc gtccatgttt ccggccacac 12540  
gttgctcag gatggccatc agactgcca gtgcggcggt tctcagttc ggcaactagc 12600  
ttttccagac gtggaacgat gtcgtgcggg ctgggacttt cctcgatttc acggcgatc 12660  
agcgcgcgt tggcagcgaa cgacggctcg tcgagcaggc gggccagctg acctcgcacg 12720  
tcgtcttcgg taaacgtcgc ggggtcgagg accagaccgg ctccccgac ggctaggagc 12780  
tccgcctac gagattcgtc ccagaaggtc ccaggagaaa tcaactgcgg tacgccgttg 12840  
accgtggcgg tttcctgcgt cgtcgtcgag ccattgatga tgatcactga acacgactcc 12900

agcagttcgt	tgagcgggtac	gtattcgtgg	acccggacgt	tcgagggcaa	ctcccccatc	12960
tcccgtaactt	cgtcgtcaga	caaggtggcg	atcacctcga	cgtcgagcct	ggccgcgccc	13020
cgcaacaacc	tttccaccat	tgtttgttcc	tggtcacttt	gctccccgta	ctgttcggtc	13080
accctgctca	gcccgcgctt	ggtcagcccc	cgcgtgatgc	agacgcgcgg	cttcgtcggg	13140
cgttcgcgca	accactccgg	caacaccgcc	ggaccgttgt	acggcacgaa	acgcatcgag	13200
atgtagtcca	agtccacagg	caggcgcata	caggatggaa	tcggatctat	ggtcgcttgg	13260
cccgtcacga	tctcttcata	gaacgtggct	ccgaacttgg	cgagcttcgc	tcccagccac	13320
gtcccagagcg	ggtcgacgcg	ctgtctcagge	ggcttcgatt	cctgggtattc	gaggaaaccg	13380
gaccgcagcc	accccgacac	atcgagggcg	acgagcatcc	gcacgtgtcg	tgccgcgagc	13440
gctcgcgcca	caactggccc	cgaacacacc	atagcgtccc	acacaacgag	atcaggctgc	13500
catttctcgg	cgaaccccat	gagatcgtec	agcgatcggt	catccacaag	gtggagttgc	13560
tccacggcgt	ccgtgtctct	gccggagttg	attgacagca	gctcgtcgaa	gagttccgga	13620
tgccgcccct	tctcgtcgaa	cgagaccccc	ctgccgagaa	cgagttcgtt	cttggccgcc	13680
aaggagatga	ggtcgagctc	gtcgcgcgacg	ggaaccgcgg	tgagtctctc	tccggtgacc	13740
atcgacacca	tattcgggca	gatggcgaca	cggacctcgt	gccccgcggc	acgcaacgcc	13800
cacgccaaacg	gcaccaggtt	gaagaagtgc	gaactcgcgg	gcagcggggg	gaacaggaca	13860
cgcatgaact	ctccgatcgc	aattgaacac	ccgggaagca	tgccaagaat	cacagaaatc	13920
tctgatatcc	ccggggaaac	gccgctttcg	caagccaaat	cttaggcctt	ccaggtgatg	13980
gtagcgatct	tgacaagcgc	gagcaggtcg	ttcccgttag	cctgggctct	accgagtcgg	14040
gtgtgcccgg	tagatcgagg	atttctgagt	caatgagcgc	ttctccttgc	tccgctgtcc	14100
tgatgtcccg	caccgcatac	aaccagggca	ggaaggtgta	aggcgcgcgag	acagcacact	14160
gtcccgtctg	gacgtcataa	cgcgattcgc	cacgggcata	gctcatctcc	tgaaggcaag	14220
gcgcgaagac	tgatcgtcgc	ctgcatgcag	agccggaaaa	ccagagcgct	ggggaaaggg	14280
cgccgcagag	tgacttcgtg	tgacgacact	tgcgctaccg	ctactgagat	gggtgccggat	14340
gccaaaggacc	ggatattggc	atccgtacgc	gattaccacc	gcgaacagga	atccccgacc	14400
ttcgtggctg	gatcgacgcc	gatccggcca	tcgggcgcgg	tgctcgacga	ggacgaccgg	14460
gtggcactgg	tggaagccgc	gctggagctc	cggatcgccg	cgggcgggaa	tgacggcgga	14520
ttcgagagcg	agttcgcccc	cttcttcggc	ctccgcgaag	ctcatctcgt	caactccggt	14580
tcgtcggcca	atctcctggc	actgagttcg	cttacctccc	ccaaactcgg	cgaggcacga	14640
ctgcggcccc	gcgacgaagt	gatcactgcg	gcggtcggct	tccccacgac	gatcaatccg	14700
gcgggtccaaa	acggactcgt	cccgggtattc	gtcgacgtgg	aactgggcac	ctacaacgca	14760
acgccagacc	gcatcaaggc	cgccgtcacg	gaacggacgc	gagccatcat	gctggcgcac	14820
accctgggca	accccttcgc	cgctgacgaa	atcgcggaga	tcgcaaaaga	acacgagctg	14880
ttcctcgtcg	aagacaactg	tgatgcgggtg	ggatccacct	accggggacg	gctgaccgga	14940
accttcggcg	acctgacaac	ggtcagcttc	tatcctgccc	atcacatcac	cagcggcgag	15000
gggtggctcg	tggtgaccgg	cagcctggaa	ttggctcgca	tcatcgagtc	gctgcgtgac	15060
tggggacggg	attgctgggtg	cgagcccggc	gtggacaaca	cctgccgcaa	gaggttcgac	15120
taccacctcg	gtacccttcc	accgggctac	gaccacaagt	acacgtttct	ccacgtcggg	15180
tacaacctca	agaccaccga	cctgcaggcc	gcacttgccg	tgagccagtt	gagcaagatt	15240
tccgcattcg	ggtcggcacg	ccgccgtaac	tgccgacggg	tgccgcaagg	gctgtccggg	15300
ttgccggggc	tgctgctgcc	ggtagccaca	ccgcacagcg	acccgagctg	gttcgggttt	15360
gcgatcacca	tcagtgcgga	cgcggggttc	acccgtgccc	ccctgggtgaa	cttccctggaa	15420
tcccgcaaca	tcggcaccgc	actgctgttc	ggcggttaaca	tcacccggga	cccgcccttc	15480
gagcaggtgc	ggtaaccggat	cgcgcgacgc	ctcaccaaca	gcgacatcgt	caccgaccga	15540
accttctggg	tcggcgtcta	cccaggcata	acggaccaaa	tgatcgacta	cgtcgtcgaa	15600
tcaatcgctg	aattcgtggc	caagagttcc	tagcatccag	catggcggca	tctcggagga	15660
tttcagcaac	gtgatcaacc	tgaccagcc	gatcctcggc	accgaagaac	tcgacgcgat	15720
cgcgagggtg	ttcgccctcca	actggatcgg	gctcggggccg	cgcacccgga	cgttcgaggc	15780

cgaattcgcc caccacctgg gaggtagatcc cgaacagggtc gtgttctctca actcgggggac 15840  
 tgccgcgctg ttccttaccg tgcagggtgct cgacctcgcc ccaggcgacg acgtgggtact 15900  
 tccttcgata agcttcgtgg cggcggccaa cgccatcgca tcctccggtg cccgcccgtt 15960  
 gttctgcgac gtgcaccccc ggacgttgaa cccacgctg gatgatgtgg cgagggccat 16020  
 cacgcggcg accaaggccg tattgtctgt ccactatgga ggatcgccgg gagaagtcac 16080  
 cgcgatcgcc gatttctgcc gtgaaaagg cctcatgctc atcgaggact ccgcctgcgc 16140  
 ggtggcatcg tccgtgcacg gcaccgcttg cggaaacctt ggtgacctgg ccacgtggag 16200  
 tttcgatgcy atgaagatcc tggtcaccgg ggatgggggc atgttctacg cggcggatcc 16260  
 ggagctggcg caccgcgcaa gacgactcgc ctaccacggt cttgagcaga tgagcggatt 16320  
 cgattcggcc aagtcttcca accgctggtg ggatattcgc gtcgaagaca tcggccagcg 16380  
 gctgatcggg aacgacatga cggcagcgct tggcagcggt cagctgcgca aactgccaga 16440  
 attcatcaac aggcgtagag aaatcgctac gcagtacgac cggttgcttt ccgatgtgcc 16500  
 ggggtgctct ctaccgcca cgctaccgga tgggcacgtc tcgtcacact acttctactg 16560  
 ggtccagctg gtcgggaga tccgcgacca ggtggcgag caaatgctgg aacgcggcat 16620  
 ctacacgagc taccgctacc cggccctgca caaggctccc atctaccgcy cggactgcaa 16680  
 gctgccttct gcggagcacg cctgcgcgag aacactcctg ctaccactgc acccaagcct 16740  
 tgacgacgcc gaggtgcgca cgggtggctga cgagttccag aaggccgctg aacaccacat 16800  
 cagccaaaga tcaccactcc gaaagtgagg atgtcgcgcy tgagcgacac attcgagaa 16860  
 acctcctcgg tatacagccc agatcatgcc gacatctacg acgcatcca ctccgcgcyt 16920  
 ggccgggact gggcagccga ggccggggaa gtagtccagc tcgtacgcac caggctgccc 16980  
 gaagcacagt ccctactcga cgtcgccgtt gggaccgggg cgcacctaga gcgattccgt 17040  
 gccgaatacy cgaaggtcgc ggggcttgaa ctgtccgatg cgatgcggga gatcgcgatc 17100  
 agacgagtc ctgaggtacc gattcacatc ggtgacatcc gcgatttcga cctcggcgag 17160  
 ccattcgacg tcatcacctg cctgtgcttt accgcggctt acatgcggac cgttgacgac 17220  
 ctgcgacgcy tgacgcggaa catggcccgg cacctggccc ctggtggagt cgcggtcatc 17280  
 gaacctggt ggtttcccga caagttcatc gacgggttcg tcaccggagc cgtcgcgac 17340  
 cacggcgagc ggggtgatcag ccggtatcgc cactcggtcc tggagggccg tacgagccgg 17400  
 atgaccgtcc gctacacagt cgcgaacccc accgggatcc gggacttcac agagttcgaa 17460  
 atcctctcgc tgttactga ggacgagtac accgcgcgc tcgaagacgc agggatccgc 17520  
 gcggaatacc ttctggagc accgaacggc cgaggccctg tcgtcggaat ccgcaactga 17580  
 gcccggaaca acgacgcaag acctggctgg gcaggccccg ggactcactc gaccgataag 17640  
 ccgacgtgac ccgggatcac ttgatcgctc tctgatgccg gcaagttttg accgtcgtga 17700  
 tctaggtgcc aaccaggtcg acgaagccct gggccctagc cgcgatcgct ggcaaggctc 17760  
 ccgaggacga tccgcctcga aatctgcgac cacctcgccg cagctgttcc ggcagacgtt 17820  
 gtgatccgag ccgacgagaa ccactcggtt cagcccggca cctgcggctg ggaggtgttc 17880  
 ggctcctgag cagcccaaaa gcaccggcca tcggcgaaga actcgaccgg ctcccgcgcc 17940  
 ggatgcgcct acccgacctc cgcgaagccg caccgcctat gctggcaacc gccacgcct 18000  
 gacctaaagg gccggtgagc caggtctggc ccagtgtcgc caggcaaccg cggcgcccg 18060  
 gcggtaacgg cgcgggaagt acaggttgct aatgaccgag acggcgaagg ccgacatccg 18120  
 taccacaggg aacaggtgtt gttcggcgta cccggtcgca gcgagcctct ccgacttcga 18180  
 ttgcaccatc gccgactgag cgcctacca gggcacattc ccggcgtcgt cgaacatgcc 18240  
 tccggtcggg ccgtcgtctg gcagcgtggc gagccgaatg gcgatcctgg caccgtctgc 18300  
 cggcgtgctg gttccgtgga agccgttaag gtcggtcgcy acgtagccgg ggcaggcgtt 18360  
 gttgattttg atgttggtat cgtgagttc cttggcgta tggatgggtga tcgcttgag 18420  
 gaacgtcttc gacggtgagt aggtccgct gatcccgccg aggtcgacgc ccggcgtgg 18480  
 ttgcaaggtc aggaagcga cgtggctgga ctggttgacg atccgcggg gctcggagcy 18540  
 gcgtagcaac ggcagcatag cgttggtaac ccgaacgac ccgatcacgt tggctccac 18600  
 caccgcccgg aggtcgcgg gtgtgacggt cgagggtcc tcgggcatg cccggcgat 18660

gccggcggtta ttaaccagca catcgagccg gccggcgcgcg tcctcgagca gagccgcagc 18720  
 agccgcgacg ctgcgctcgt ctgtcacgtc cagggatacc gcgaacgcat cgacgccgtc 18780  
 cgcacgcaat ttccgccacgg catcctcccc gcgctggtgg tcccgtgccc cgattccgac 18840  
 gctccacccc agcgcgccga gcccggccgc gatctcgtat ccgattccct tgtttgcgcc 18900  
 ggtgaccagt gcaatcgtct gttcgctcat gctttctagc gtggcccga accggcgggg 18960  
 aaccaacacc atctaggtct cacaccgata cccagctgat atgaatcagc ggtaggctcg 19020  
 acggatggag acgcgggagt tgcggtactt cgttgacgtc gccgaggagt tgcacttcgg 19080  
 cccggccgcc cagcgccctgg gcatcgccca gccgcgctg tcgcggaaga tcgccagct 19140  
 cgagcaacga ctcgagtcg tgttgctgca acgcaccagc cgcaaagtct cgctcaccga 19200  
 agccggggca atgctgctga ccgaaggccg ggcgatcctc ggcgcgctgg cagcagccga 19260  
 gcgacgcacc cagcgtgccg cgacgagcca gccctcgcta gtccctggctg ccaaggccgg 19320  
 cgctccgggt gagctgctgg cgaagtgtct cgacgcgtac gccgcgagc cgggagccgt 19380  
 ggccgtcgac ctgctgctct gcgaatccca gcccagaaa acgctgcatg acggccgggc 19440  
 cgacgtggcg ctgttgcatc aacccttcga cccgacggcc gaactcgaca tcgaaattct 19500  
 gaacaccgag caacaagtcg ccattcttcc gacctcgcat ccgcttgcca gcgagcccca 19560  
 tgtacggatg gcggatgtca gctcactgcc ggatctcccg cttgcgcgct ggcccggccc 19620  
 cgacggcgctc tatccagatg gcccggcgt ggaagtacgc aaccagacgc aactgttcca 19680  
 aatgatcgca ctcgcccgca ctaccgtggt catgcccga tccagtcgcg tcaacctgct 19740  
 cgaaggccctc gccgcgctac cggttctaga cgcgcgggac gtgacgacag tcatcgccctg 19800  
 gccgcccac agccgctccc gagcactcgc cggcttggtc cgcgtggcca cactcctcta 19860  
 aatcttcgct cctcaattg cccgcagcta ggacgaatgc cgtgcgactg ccggaacga 19920  
 ccgcgcgcaa cacctcatgt gcctcagcgg gagaagacgc tatgcggatg ccgcagagac 19980  
 ctacgggccg tcccgtcgcc aggagccagc ggtcccgttc gagctggtag gtttcggcga 20040  
 tctgctccgt gacgcggtcg atgtagccga agctggttgc gagcattcgt tgcgtggtcg 20100  
 ctgcggatac ggccctcgtcg gcgcattggc ggcggatctc gtcgaggcat cactacagga 20160  
 agcggaaagt cccgatgcgg ttggccctga tcagcgcggt gatggaaatg tttcgccagt 20220  
 cgtcgcggat actccaacac cgcagcagga cccggcgaat atcccaacgg aatctcgaat 20280  
 tcgagaacgt gcggcaatgt ggcgatattt tcctcaacac tcgtccgtag aatgctgagg 20340  
 acgatctcat cggctcgcaa ctgcggagcc tcggtaatca ggtgcgtgta gacgtctcgg 20400  
 ctgacttccg agacgcacga cgcgacgctc gatgccacgg ccgcctgtag cttcgaggaa 20460  
 tcgaagctgt ccgtggccat gacacgacgg tactgcggcg ggagcagaag cagcgcttcc 20520  
 tgcgccgatg gcgcgcctgg agaagatcag caccttcggc ctgaccgagc cgggcgcacg 20580  
 gcaccgacgt ggtccagctg gagctcggag acacagcgtg gcggtgttca gcgccgccag 20640  
 gtcccgacg cagcgctttg gcgtgcggcg cttcgggaga gctgggccat tacaggttgt 20700  
 gcgagagcag ctggctcagc acctcgtcga cgttcatcca ccggtgttca tcaccaggtt 20760  
 cgaccacgtc gtaggtgtcg ttcaagaagt ccgcaagctg gtcagcattc acctcgaagg 20820  
 aggctgccc cgacggcgag ctcatctcga tcacgaccaa ccccggaata ccccgtcgag 20880  
 ggccgatccg cacatcgctt tcgcctgcct cggcgatcag cccgtcggcc agcaggtcgc 20940  
 gggcgatcac ccagtcacac tgaccgtccg tgcctacgtt gagcttcatg gagatctcgt 21000  
 acggattgcg gctgtcgtac cgcagctcca cgcgactgg cgcgatcaca cccgccgggg 21060  
 ctaccaggtt gaagatcatg gtggcgagca ccgtgcgctg atcgtttgc atgatgtcca 21120  
 cgctcccaag ccttggttcc ccataccctc ataacgaagc atgcccaccc cgggttcata 21180  
 ccttcttcac gaagtcttca gacctggtg gcagtcctcc tcgggctccc agcgaggaa 21240  
 cttcactcac actgaaccca cgggactctg ggcccggaac tccgaagct gacccaaggt 21300  
 gcaaagcaac cggcgagtag cttttcgtg cgccagtcaa tcaatgacgg ctccgggaat 21360  
 gtctcgtgac catgctaggg cgtactcgcc gaaaaatgg atggatacgg gctgaccagc 21420  
 aggaacgcgg tcagacggtc tggctcctcg gaaagaatcg ggcgggtttg tctggcgcg 21480  
 gcgtcaggac tggccaagg cccgttcgat cgcgcggcg gcccgccgg cggccgcccg 21540



gtctgtctgg agctgcatgg ccgctgttcag tgccagcccg gtggcgacga tctcgaacag 21600  
 cgcttggtcg atgtcgaacc cggcgggag ctcgccgttc tccaccgccc cggtcagggtc 21660  
 cggccgcagc tgtccccgcc agcgcgacca cacctcggcc accgcgtcgc ggacccggcc 21720  
 gggcgggccg tctactcgg tgagcgccgc ggtcatcagg cagccgcccg gcagcagtgg 21780  
 cgcttctagg tagccacgg cgttgggcga caccgcgcgc agccgcccga gaccggggcg 21840  
 ctcggccagt gcgggctcga ccaccgggtg ccagaagtcc acgaatgcct tgtccagcgt 21900  
 ggagatctgc agtgtctcct tgggtgccga gtgcttgtgc accccggact tgcctatctc 21960  
 cagttcctcg gcgagccggc cgatggtgat gcggtccagc ccttcctcgg atgcgatctc 22020  
 ggcggcacgg tcgaggatcc ggctcctggg ggctgtgtct tcggccgctg agcgtcgcgg 22080  
 tgacatgtga cgaggatagc gtacgcgcgt tcgctattga tttagtgaac gtgcgtacgc 22140  
 taaattccgc ttcaccacg gaccgaagga gtgcgagatg cgggtgagac gactgggctg 22200  
 ggccggactg gaaatcgagg cgggaggcaa gcgactggtg atcgactacg tacgggacct 22260  
 ctcaccgctg ttcacggcgt ggaaaccggc cgagaggctg gcggtgccga gcgggacggt 22320  
 caccgcgcga ctggtcacc accctgcacc ggaccacacc gacgcggccg cgctcgcggc 22380  
 cgcgtgaca ccgggggac cgggtgctccg accgcgcgcc ggccacggcg acgacgtgga 22440  
 caacgtgacg aactgcccg ccgagcgca gctgacactg caccgactgg ctgccagggt 22500  
 cgtggatgcc tgggtccacc gcgacctcgg gccgttccgc gtcaccgagg tccccgcgt 22560  
 cgacgggctg ggcgaccgc agctgaactg ggtggtgcag gccgacggg agcgggtctt 22620  
 ccacggggc gacacgatgt tccacggcta ctggtggctc atcgcgccgc ggttcagccc 22680  
 gttcgacgcg gtgttccctgc ccgccaacgg cgcggtggtc gacgcgccac acctccagcc 22740  
 gccgagccc ctgcccgcg cgatggacc gaggcaggcc gccgcagccg cagagatcct 22800  
 cgacgcccg tacgcagtgc cgatgacta cgaggcgga cagccggaca agatcgcggg 22860  
 ctacgtcgag gtctccgacc cggagaatga gttccgcag caccgaggac accgcgcaca 22920  
 cgtgctgccc atcgaggat ggctggacct ggccacttga ggcggaact ccgccgcgcc 22980  
 ggcgggctca ggtcgcaaca gaacctagg ccgatggatg gtccggaatc gcgcgggaat 23040  
 cggttgctcc aagacactct ccggtcgcgg cggcgtcaat cacaagattc agcatccatt 23100  
 actgctgga accacgatca aatgcgacac tccgggttga gccgaacca cgggtccagt 23160  
 cctggctcctg ggcgggacgg accttggtg tccgcgcgag gaacctcac ttgcgtgaa 23220  
 cccacggcg ccgaagcgtc ccggtagcag acgcgggaaa ccagcgctta gctctggaac 23280  
 gctggctgcg gaccgggccc agaggcgagc gctcgggtgg tgctgagttc ggggacgtac 23340  
 ttgtagatcg tcgagcgga gacgcaccgc acgcttgagc ttgaacggcg cagcacgca 23400  
 cccgcccag ataccggtta gcgtggtcag atcaacactg cccgttgtgc accattagcc 23460  
 gcatatgtcc ggttaccatc gattcgatgc ggatccggtc cgtgtcaaag ctctctacgt 23520  
 cgacgggtac ctgaatgcc cgctcaacga cgcgtcatcc gggtcagccc ttctcgagca 23580  
 ggcgagcgcg ctggcgggcg agttgggca ccaatcaggg gcggcctacg tcgcgcagat 23640  
 ctcggggctc accgcctgt tcaacgaaga tccggcgcg gcgagcgcg tcttcgaaga 23700  
 cgtctcgcgc aagcatcggc tgattggcga tcacgctgcc gccgcctacg atcagatoga 23760  
 gctcgcgttg gcagtcgcgt tccttggtga tcatgagcgc gccgctgcc tgttcgcgg 23820  
 gtgcctgcc gtcaaggaga cctacagcga acagtggatg agatcgcttg ccctgtgggc 23880  
 gaagggaata gcagacctac tagcgggcaa ttaccggcaa gccaatgctg cggaactgga 23940  
 gagtcttcgc ctgcggctcg gttccatga gcagtttcag atcgcgctgt gcgttgagat 24000  
 tctggcctgc atagccagcg cagacggtga cccggcacgg gctgcgacct tgttcggcgt 24060  
 gtccaaaacg atcaaaaaga acgtcgatgc gtcgctcgcc ggccacaaac atgttgcgcg 24120  
 cctacacgac cactacgagg cagtagcac tagctcactc ggcgacgaga cgtttcagaa 24180  
 aggcgttcac caccgcacac agctcgactt cgacgaggcc atcgcatgga tacggaatcc 24240  
 gggcaaagcc ggggaaccag caagctccag gacggaagaa acgcacgcgt cgggtgcttac 24300  
 tccgcgcgag cgagaggtag cagcgtggt cgcacagggc accaccaaca aagaaatcgc 24360  
 cgccatcatg gtgatctccc gccgcaccgc tgaggccac gtcgagcaca tctgacca 24420



actcggcttc accactcgga cgcaaatcgc gacctgggtc attcaacaga agaccgcgcg 24480  
 acgatgcgac tgacaaccgc cggttttcgg attggcgtgg accagtagaa gctgccccgt 24540  
 ctgcgagacg gcccctctcc acagcctctt ctccacagtc gatcaacgac cgaacgggac 24600  
 cgcttgagga cggcatccga tgcccaccgc gcgagagga gcacttgcc atgcgcacac 24660  
 tgatcgagaa cgctgccgtg gtgacgatgg acgacgttct cggcgacttc gaacgagccg 24720  
 acatcctcgt cgacaacggt gtcacgtcg agatcgggcc gaacctcaac gttggcgaca 24780  
 ccgagcgaat cgacgcctcg tcgatgatcg ccatgccggg catggtcgac acccatcgac 24840  
 acacctggca aaccgggctc cggggcattc tggcggacgg gaacatcctc gactacctga 24900  
 ggggattccg gctgcagatg gcgacgaagt accgcccga ggacatgtat gcgggcaact 24960  
 acctcgggtg cctcgattgc cttaactcgg gggtgaccac gctggtcgat tactgccaca 25020  
 acatcgtcac cggcgagcac gcgcacgcc cggtcgctgg actccgcgac gcgggcgtgc 25080  
 gcgcgctgta cggccacggc ctctgccc tcaacgtcgaa cacgtggtcc gagaccaaag 25140  
 gcgggctgga cgagtcggtg gaagaaggcg acttcgcccg gcgggcgcgcg ctggcccggg 25200  
 aaatccgcgc ccagtacttc ccgagcgagc aacaattgct gcgcttcggc atcgccccgc 25260  
 aggaactagc catcgctccc ttcggggacg tcaagcagga attcgagctg gcccgcgagc 25320  
 tcggcgcacg catcacgttc cactgcaacc aagtcatcgc gcgcaacctg ttcaaggaca 25380  
 tcgaggcgct gcactccac aacatgctcg gcagcgacct tctcttggtg caggggacat 25440  
 tcagcactga gaacgaatgg cacctgctgc ggggtaccgg caccatgatc tcggtttgcg 25500  
 ccgagaccga gatgcagatg ggcattggat tcccggtcac tcgtgaggcc accgagaaca 25560  
 cgccggggcc aagcctcggc atcgactgca cgagcagtac ctgcggcgac atgatttcgc 25620  
 acgcacggct cgtcctacag gtgaccgat ggaggagcga tcaagaagac tacgcgcggt 25680  
 cgacactgcc caggtcatg cgggtgaaaa cgcgcgacgc cttacggtgg ctcacggtca 25740  
 acggtgcccg cgccgcgggc gtcgacgacg tgacaggaac cctggcaccg ggcaagcggg 25800  
 ccgacatcgt attgctggac atgtcaggca tcagccaggc cggttggaac cgccacgacc 25860  
 cttgtggcgc gatcattgag cagaccaact cgggcaacgt gcacacggtg ctggtcgatg 25920  
 ggcgggtggt caagcggaac ggccggtggt tgcacgtcga cgtcaacgga gcgcttgcaa 25980  
 cgctcgcgga gtcgcacggt catctctacg accaaatggc ccagcacggc ggattcatcc 26040  
 cgcaacctcc agccgagcta ccggtgttca accgctgaga accaagccag ctgtggctgg 26100  
 cgcgccattc gctgcgtgag gtgaccggc tggccgggct cgaccggaaa accgtcagac 26160  
 gctgctgtgt gacctgcacg gacgaggtta tcggcatagg caaggccagt cgcagcagca 26220  
 cggtcgctca tcatgccgcg cccaccgtca gccgcgtgcg agcgccgatt cgagcaacca 26280  
 cggccggggc cctggaatag ccgcttccac ctgcacggag caggtattcg cctcgtggac 26340  
 cgatgatcgt ccccgatgaac gcgtggggct gccgtgcggg aacgcgccgt cttccggggg 26400  
 gaaggcgggc tgcggaacc ggcgttgact gacctcgt gcgcattcga gcatcagcgg 26460  
 ccacatcaat ggggagggcc cgggtgacgac aaagaccgtc aggtgtggtg gcaggacgcg 26520  
 gtccgaacgc accgtgaccg ttgaacggtc ggccgagctg tcgtttcttg ccgcagcaac 26580  
 agcaactcga cgaacgcgcc ggcgatcttg ccgcgcggg gccaccaacc gcgaactcat 26640  
 gaccgggctc aacacatcaa cccagccccg ataacagggt tcgagccctg cctgtcgact 26700  
 cgtggcaatc gccctaaagg agggccagct gtggttacgg atcgtgttcg cgtgctcctt 26760  
 tgtgaagcgg gcgtgagtga gttcatagag cccgcgcgcg gttcgtgagt aagttcttga 26820  
 acatggagaa cggcggaccc agccggacag cgctcgctac cgcgtacgcc cgcgcctatc 26880  
 accagatcgc tgaccgacca cagatcttca ccgacccctt ggcggtacgt ctgctcggcg 26940  
 tcaccgcaga ggaactgacc aaattggcca cgcccacaac cgatcacctc ggcgccgggg 27000  
 caatcgatca gcccgccgc ctgttcttcg ccgcccgtgc ccgcttcgcc gaggacgccg 27060  
 tggccgcggc catcgccgat ggtgtacggc aggtcgtgat cctcggcgcg ggcttgaca 27120  
 ccttcgccta tcgcaatccg cgcccgact tgcgcgtctt cgaggtcgac caccgcgcca 27180  
 cccaagcatg gaaacgcgaa cgcttgcta gcgcaggat cgaccgccct gagacgttga 27240  
 ccttcgtacc ggtcgatttc gaaaccagc cgctggcaac tggattggaa tccgccggat 27300

tcaaaccggac ggaccgcgga gatttcgtgt ggctcggcgt tgtcttctat ctgaccgcga 27360  
acgccgccca cgccaccctc gaatacatcg ctggtcaagc ccagccggtc gaggtggtct 27420  
tcgactacct gcagcctgcg gccaccgacg aggaccgtgc gcaccagcgg gcacgtgccg 27480  
atcggtctggc cggcgccggc gaaccttgt tcagctactt cacaaccgac gacatcgccg 27540  
cacagcttca cgccctcggg ttcaccggca tcgaagacca ctccgccccg gacctcatcg 27600  
ccggtacct ccacggatcc gcaggtttcg agggcggaacc accccgcgca ctgcgcgca 27660  
gccgcatact gcacgcgagc cgctgaggtt tgcccgccgc gcctacttgg accacatctg 27720  
tcgcgacggc tagcatctcc cacatcgctc tctgaactgc acaaactcaa aaaatcgcg 27780  
gagcgacgc gggtagggcc ccggcgaaaca agccctagcg atacatcact ttgaacaagg 27840  
tcggtctcg ataggtgcc aatgcgccgt tgcgtttcgc gcgaaccgcg cggtcgactt 27900  
ggctcgtggcg tagtcaagga accgctgcag ctcatcgcg gtgaacgggc ggtcgtccg 27960  
atcgccctca taggactgga gatggcgag tgagttccac tcgtgcacat tcgccaccg 28020  
gtactcgctc gggccgaacg acgctcgca cgccagcgcc cagccatacc ggctgtcaca 28080  
gaggaattcg ttgaacagcc tgagactgcc ttggtaggcg cggattgtgg atggcgcgag 28140  
atgccgttct ccggtcgggt attgcgacca ttcgtccaca tgcgacggtg tccaccgcca 28200  
cggatactcg ttggtgaact ccgcgaaccg gcgcaccagc cgctcccgca ccccgccag 28260  
gtgcgggaac attgccttga tcaaaccgca gtccagacga tcatgatcaa cgagtgtgat 28320  
caccagcagt cactggagcc ggcgttgct ccacacggtt ggccaaccg agcaggtggg 28380  
tgcacagcaa cgttgttgaa gcacaacgcc tgactggttc aggtccggcg gcagccgta 28440  
ttccagccgt ctccggcgcg ggacgttcgt gttcagagat aaaatcgtgt tgcccgttg 28500  
gtcgttttg gcggttttga gtccgtgcga ggccgattat gcggcatggg gcctgtgttg 28560  
aggaggtgcc gtggttacta cgcagcgga accggcggt gctgggggaa cgacggtgag 28620  
tctgccgggc gtaacggccg agttccaccg tcccgatcac tacattcca ctcgatga 28680  
tcttgctagc gcggtgag ccgttcgggg cgttctcccg gaccatcaga cgatggcctt 28740  
ctatcggtg cttggagcgc tgacctgtt ggggttgatg gactggccg tggcgcggc 28800  
gattggcggt ggtgtcgcg tcgcgcagcg tgccggcgga ccggcgctga agtgatccg 28860  
tttgtagcgg gttccgggct aggcgtcgct gagcgcgccg ccggcgcgct caggccaggt 28920  
gaacctgggt aggtgttgcg ttgcctgcgt ccggcgtagc tgcatgctcg gagccgtttc 28980  
gaagaaccg agcacacag cggccttgaa ttcaccaccg ctgggactac tgggtgtgctg 29040  
gcgggtcgaa cagcgtggcg aagccgtcg tgttgtgat cagctcgag gccaccgaga 29100  
ggacggatgt gcggtgttt cgaggttcgc tttccgctcg tcggtctac gctcgtgact 29160  
cacggtatct gtgcgcgcg gaggtagtgg tggctaggac ggcgcggaac acaacgcctg 29220  
gcgcaggcgc gcggcaaca gggcgcggt tgcgagccgc ggcggtcagc ctgcccgtcg 29280  
tggcacagtt tcgtcgggtc gatgttcgca tgccgggtcg ttgggaggac gcggttcgcg 29340  
cggtacgggt cggttggggg ctcgaatggc cggttgcggg ggcgaccggc gcgggggtgt 29400  
cacggcgctc ggtcggcttt gcgtcgga gctgagtgt ttcgatgctg cgtaccttg 29460  
gcgcgttggc ctgcagatc gtggcgccg tggtggctgt cgcgcagagc atcgcgggct 29520  
tggggcgcg gcggcgggcg tggctgctcg agggcggtgc gtacgtcgag gtgcgcggtg 29580  
ttgaccagcc gggggtgag gaggtggcg agcagatcg ccggcggttg gagaaggtgc 29640  
cgggtgttgc gtgggtcgag atcaacgctg tcgtggggcg ggtggtcatc gggcacgac 29700  
cgggcctggt tgggtcgct gagctggtg gtgtggtgga ggacgcggag gcggacaccg 29760  
gtttggacga ggaaccgttc gctgcggtc gtgtgagcca cccggacaat ccggtcaagg 29820  
tgatcgctga ggcgacggt ctgggtgca acttgaccg cgtgtgatc gcggtggggg 29880  
cccgatgct gccgtgcc gtgtgccgc cgtctgttcc ggcgctggtg tcgaccggtg 29940  
agtcggcctc ctgggtccg gggcgctg aaacgcggat cgggaagacc gccactgacc 30000  
tgttgttcg gtttagcagt gcgttgggga acacgcttgc gcggcgctcg gcgagtttg 30060  
tgggcggaatc ggggtaccg ttctgctgt tgccggagac gcaggcgcg ccggcgagct 30120  
gggcgcggtg ggaagccgga acggggcg atcctgccag tcaccgtagt gcgcgggtg 30180

ggggttcggcc gcgtcctgtg ccgttgcccg ccggcccggg cgagcgcgcc gcgaacgggt 30240  
 cggctctggg tgggccggcg gcgttcgcgg gcatgctcgg ggtgacgcgg agttttcagc 30300  
 gtgcgcaggc gatggtggtc gcgggcgcgc cgcgcgcggc gcaggtgggg cggggcgcgt 30360  
 tcgcggcgca gctcgggtac gggctgtcgg agcggacgac tctggtgttc gatgcccggg 30420  
 tgttgccggc gttggaccgg gtcgacacgg tggatgatga cgcggacgtt ctgcgcaccg 30480  
 gtgctcgggt ggtgcatgac gtgattcccg cggacggggg gcgcgcggtc gtcgagttgt 30540  
 gggaacgtgc gcatgaactc gtgggtgtga ccaagggggg cggcgcgcag cgtgatgggt 30600  
 ggtcgtgta ctcggtcgct cggtcgcggg tgccccaga cgtgcgcgag gtggccggtc 30660  
 ggggtatcga ggtgttcgcg ctgcgccacg attcgcgcac ggccgcgtgg gttctggctg 30720  
 ccgaggagct ggatccgctc gctgaggagt tggtcgcggc agctcgcgag gtcgggtcgg 30780  
 tggttctcgc gggatgatgg ggccgggttg cggggcggat cgggtgtgat cgcgtggtcc 30840  
 cgggtggggg gcggctggct tcgggtggtc gggagatgca gcgggagggc agcgtggtcg 30900  
 gtgtggtgtc cgcgcggggc cgggcggcac tggcggctgc ggatgtcggg cttggtgtga 30960  
 cgcgtggtgg tgaggttccg tggggtgcgc acctgatctc gggcgcgagc ctgggggagt 31020  
 cgtgtgcgtt gttgagtacg gtcccagacc cgcgttatgt cagcgtgcgt agcgcgcagc 31080  
 tgctgttgat cgggttcggc agcggggcgg tgttcgggcg gttcgggcct gccatcggga 31140  
 gccaggctcc tgccggcggt ccggtggccg tgtgtgcggc aatcgcgctc ggcgctggga 31200  
 cgtggtgggg gatgcaggcc gcgtggcgct cggttcccgc cgcggtgccg cgcactgcct 31260  
 ggcattcgat gggccccggc atggtgctgg agctgctcgg cagttcatcg cagggtctga 31320  
 ccgatgatgc ggcgcacgag cgtttccaag gtcaggacga ggggacgggc cttcgccgga 31380  
 tgggtgtgtt gcgggcgtcg ctggaggagc tggcgagccc gttgacgccg gcgttgccgg 31440  
 gtggggccgc gatctcggcg agtatcgggt ccgttgcgga cgcgacgac atcatgggcg 31500  
 tgctggggat gaacgcgttg atcggcggcg tgcagcgcac caccgcggac cgcgctttga 31560  
 acaggcttct ggaggccagc gttgttcggg ttgtggtgcg ccgtaaggaa aatgtgacgc 31620  
 agttgccggc tgacgagctc gtgctcggcg atgtggtgga gttggtggcc ggggatgctg 31680  
 ttccggccga ctgccggctt ctggaggccg acgggctgga ggtcgatgag tcgagcctga 31740  
 ccggggagtc gatgttggtc accaagacgg tgacggcgac tccggcggtg ccggtcgcgg 31800  
 accggacgtg catgatctac cgggggaccg cggtcgctgc cgggcggggc gtgggtgtcg 31860  
 tggtcgcgac cggcgcgaac accgaggcag gcagcgcgat cgggccgat gcccgcgag 31920  
 gtgccgatgg tgcgcggct cgcctgcaga cgtggccaa ggtgacggtg ccgatctcg 31980  
 tgggagcggg tgggattctg ctcgtcacgg acatgctgcg taggcgtcca ttggggacat 32040  
 cgctggggcg ggcggtgagc ctggcggttg cggccgttcc cgagggattg ccgttcgttg 32100  
 ccaccgtcgc cgaactcgcg gccgcccggc ggctgtcgaa acgcggggcg ctggtgcgta 32160  
 actccgcgac gatcgaggct ttggggcggg cgcaggtgtt gtgtttcgac aagaccggca 32220  
 ccctcaccga agggcgcac gcgctgcggc aggtatccga tggcgcgtcg agccggtggc 32280  
 tcgaggaaat ctccacagtg gaacggcgag tgctcgcggg tgcttgcgg gcaagtccgt 32340  
 tgggtggagaa cggagaactg ctcccgcatc ccaccgatcg cgcggtcgtc gatggcgcgc 32400  
 gtggcgcggg agtgcacccg gccgaaggca agcagcgtg ggtaccgatc gacgagatgc 32460  
 ccttcgagcc ttcccggtgg taccacgcgg tgctcggcac ctggcgcgac ggcaatctgc 32520  
 tgagtgtgaa aggtgccccg gagatcgtgc tggcccagt cacccggttg cgcagatcag 32580  
 atggcgacgt gccgttgag gagaccgctc gtcgaggcgt cgagcgagtt gtcgaaaacc 32640  
 tcgcccggca gggctaccgg gtgctggctg tcgcgagcg ggccgcctcg gaccgcagtg 32700  
 atctcgacga gtcccgatc cgggatctgc ggctgatcgg ttctgctcg ctggccgacc 32760  
 cgggtcggcc cactgcggcc gccgcgctc cgcagctcca gcaggcaggt gtcgaggtcg 32820  
 tgatggtcac cggcgaccac ccagcaccg ccgaggcgat cgcgcgcgaa ctcgacgcgc 32880  
 tcaacgagcg tcgctcatg accggcaccg aactggacac cctcgacggc gaccaactgg 32940  
 cggcggaagt gaccaccgtc gcggtgttcg cgagggtcag cccggcgag aaggcgcgca 33000  
 tcgtcgacgc cctgcaacgc aacgaccgtg tggtcgcggg gaccggggac ggccccaacg 33060

acgccccggc gatccggttg gccgacatcg gcacgcgcgt gggcgāggcg gcaacccccg 33120  
 ccgccccgga ggccgcccgc ctctgtgtca ccgatgaccg catcgagacc atcgtcgacg 33180  
 ccattgtgga aggccgcgcc atgtgggcat cgggtgcgtga ctctgtggcc atcctgtctg 33240  
 ggggcaacct cggcgagatc ggcttcaccg tcgtcaccgg tgtgatcaac ggcgggggca 33300  
 gcctcaacgc ccgccagttg ctgtgtgtca acctgtcac cgacatctc ccggcgatgg 33360  
 cgatcgcggt gcgaccacc ccgacgatca ctggcgagat gctgtgtggca gaaggccccg 33420  
 aagcctcgct gggcgctgcg ctgaccgcg acatctaccg ccgcgccgcg gtcaccgccg 33480  
 gggcgcccat ttgtgcgtgg atcctgggca gggcgaccgg aacctgcgc caggccaaca 33540  
 cggtcgcgct cgtggcgctg gtcgccgcc agctcggaca gacctgggtc gtgcgcggac 33600  
 ggacaccact ggtcatggcg tgtgcgcttg tctgttctgt ggcaactggcc gctgtcgtcc 33660  
 aaactccccg cctgagccaa ttcttcggag ccagcccgct actcccacac ggctggatga 33720  
 tcgccgctgg cgcaacgacc gccgccaccg tcatcagcct cttgtctcaa caatggtgaa 33780  
 cctccatgca tgagcgtcaa gatcccttca ccacagccat cgtcggggcc gactcgggac 33840  
 aacttcgagc aggaagacg ttggagcgcc gagagcggtg aacgactcgc cagcgccgaa 33900  
 agcgcccgcc ccagccgac gggcggtgga cccacaatgg gagctcggc gacgtgcgtg 33960  
 cggtcggtca tcaggatctc tgtcgtgagc ccgtccggat agatgaccgg cactgcggg 34020  
 taggactcgg cgatcactag gtctgtgtg tgcacctgga gacctgaag agccgcatcc-34080  
 ggctcggctt gaacaatcac cacatccagt cctggacgcc gccggcgtag acgagtcagc 34140  
 acctcgggga acaggatgaa ggtgacggtc tgcattggagg ccacacgcag acgttgcgag 34200  
 atctcgccac gtgagcgatt gatctgagtg cgcagcgtct ccagctcagc gatcaccgg 34260  
 tccgctgcct ccacgacgag ctacaccttc gccgtcaggc ggacgcgacg gccttcgcgc 34320  
 accagtagcg gagcggatgc cgcaccagt gccgcgcgag ctggcccgcc gcataaccgc 34380  
 cagcagtggc ctgaccagt ccaacatcgc ccacacgggc ccggtcggcg accttgatca 34440  
 gtggtacagg ctgggcaggc cagcgggcaa cgcacgccct ccccgacccc accacgatca 34500  
 acatcggccg agatcgtaac gaaccccgtc ggtgggttcg aagtgggttcg aacagttgtt 34560  
 caagctgtaa gtgcgctgcc agccatgacc gggctacaat aaaacgctag ttctaccgag 34620  
 gtgcggtgag tcgtgtctgg ggttgcggtt cgggaggctg gtcgcctccc atatgaggtg 34680  
 acgagctttg tggggcgccg tcaggcgggc acaagagtta agcgcgcgct gtccagttca 34740  
 cggttagtga ccctgactgg ttctggcggg atcgggaagt cccggctggc gctgcacgtc 34800  
 gcacacgacg tgcggcgggc cttctccgcc gggacatttc tgggtggagct ggctaacgtt 34860  
 cgcgatccca tgctggtgcc gcaggcggtc gcggctgtc ttgggatcca tgaccgttcg 34920  
 acacgtgcc cggaggcgt gctcgtcgat tacttgccg acaagcaagt cttgctgggtg 34980  
 ttggacaact gcgagcacct gctggggggc tgcagccgcc tagtcgctgt gctgctggcc 35040  
 gcggccccgc ggttgcggtt acttgcaacc agccgagaac atttagggat agtggcgga 35100  
 cagacctggc cggttcgcc actgtccgtg cctgtggca gcgattcatc aacgtcacia 35160  
 caggaaggtc gtcctacca gcgtgaagcg ttgacctgt tcgagcagcg tgccgcggcg 35220  
 gcgctaccag ggttcaccat caacgaagaa aacgaacaga tcgtggccag gttatgccgg 35280  
 aggttgagc gcattccct agcgatcgaa ctgcgcgcg tgcgcgttcg cgtgtctca 35340  
 gtcgggcaaa ttctaacgcg aatggagaac cgattccggc tgcacccac ccggcaccgc 35400  
 gccgacgatc ctgcgcatca gacctgcgc gccgcgctg actggagcta caacctgtgc 35460  
 aatgagcatg agaagcagtt gtgggcgcgc tgttcggtt tcaccgacga attgatctc 35520  
 gacgcagccg aaaacgtctg caccggcgac gggctgacc atgatgacgt cttcgccgct 35580  
 ctgcggggac tgatcgacaa gtccgtgtc acccgaccg aggagcgac gaagacgaga 35640  
 taccggatgc tggagactat ccgtcaatac gggcacgacc ggctcaccga ggcggtgccc 35700  
 gaaaacgctc tgcgctgcc gcaccgcgac tactacctgc atctcgccga gcagtccgat 35760  
 gccgagtcg cgggccccga tcaagccggt tggctcgcgc gtctgcgcgc cgagcgcgcc 35820  
 aacttctggg cagcgttgga ctactgttc accacacctg gcgaggcgtc ggttgggttg 35880  
 cggatggtga gcgcgctgtg gttctactgg gctgtggcg actatctgag ggaggccgc 35940

atgtggctgg acagggcggtt ggccctggag accgaaccca ccagccagcg tgcccgcgca 36000  
ctgtggatca ccggatggat tgcccaccgg caaggggacc gagatggggc gctggtcctg 36060  
cttaacgaaa gccgcgagct ggcccggaaa ctgggcgacg aaacggaatt tacctacgcg 36120  
actcagtttc ttggcgaccg ggagatgtgg gacaacaacc tgactcgcgc cacgcaactg 36180  
ctcgacgagg cactggcccc ccaccgcgca actggccact ggaccggccc cgccttagcc 36240  
atcttctcaa tacgagcgca ggcagcgggc ctgcttggag accttgaccg ggcgatggtc 36300  
cttctgcacg aatgccgcaa gatctgtaca gaacttggcg agcgcgtggag gctatcgtgg 36360  
gcggagtggg acatggcgat cacttggtgg gccgcgggcg acccaaccaa ggcagccgcc 36420  
tccgcgtcca ggtccctccg caataagcgg gaactcaacg acctactcgg gatcagcgac 36480  
tgcgttgaac tgctggcctg ggtcgccgcc gcagaacgca atcctgaacg cgcggcgatc 36540  
cttttcgggg cgttggataa aatgtgggaa ctgattggca ccccgctgtt cggatccgag 36600  
acgctactga cctggcggtga gcaggccaag gcacgcgccc aggaagcact cacagacacc 36660  
gaatacgaag ccgcccaggg ccaggggacg cggatgaacc gggagcagac tatcttctat 36720  
gctcttgggg agaagccacc gccggtcgaa ggcggccgccc cttcggctcc ggaaaccgaa 36780  
ccggttctca caaaacgaga acgcgaggtc gctgccctgg tcaacgctgg caaaacaaac 36840  
aaagagattg cggctgacct ggtcatttcc cagcgcacag ccgaggccca cgtcgagaac 36900  
atcctttcca agctgggctt tacctcgcgc gcacaggtca tcagttggat gaccgggcag 36960  
agataacgta gtcacctgct ctgttggccc atgcccttct cgggccagcc cggctcacat 37020  
gatccgttcg cggcccggca acaccgccc tctgtggcaag gtctgcgata ccgaaagcgg 37080  
acatccgctt cagctaatac tccgcggccc ccgcaacggc gcatgttgaa gtaggtatgg 37140  
tccatgatcg gaggtggcc attgccgatg gccatgaccg cggagttcgg gtgctgcacg 37200  
tagttgcggc cgccacata ggaccggtt gcgagcttac gcggtataga ttgagccacg 37260  
ggccagcttc gtaggccttg gctgttgccg ccgacgagta gtgagtagcc ggtgtgaacg 37320  
gtctcccagc accatcgagc caccgttcgt ccggttgta gtgccgtgc gcgaagcctt 37380  
gcacacgtcg atagcgggct cttgagccaa aaactcgtcc cgttccaaca cgagcacatc 37440  
gagtcgcgca ccagcgagcg ccccgcgac gatcagtcag cctccaccgg acccgacgat 37500  
cgccacgtct gccgatgccc ggttgtcctt gtgcatgatc tgttgtgcct aactgcgccc 37560  
agctttcgtc atcgagcctg tcccacagtc tgccgatgcc cggtcggagc agcctccagc 37620  
gccgcccgtg cgaaggccaa cgaccagacc ccgtcttggc cggaggccgg gggctcggcc 37680  
tcgccacgca cagcctcgat gaaccgtcga acacccacga cgtagaggtc ctcgcggtga 37740  
cccagatcga tcttcctgga cccggccgga ccgacgtgtg tcagcgtgcc gaccggtgac 37800  
atcgactgga taccggtgcc gatgagcgt ccttcgggtc cgaaaacctc caggcatgtg 37860  
ggggcatgtc cggtcacgaa gctttccacg aatgaagcat gcacaccgct gacgaactcc 37920  
gcgcttcggg agattgtctc atcgattccc accctcgtca tgtggccgga cgaggtatgg 37980  
gccacgaccg ttagcggatc ggcattgagt acgaaccgga gacagtcgc gtcgtgcacc 38040  
gtgaggtcga gtgcgactcc ggcgcccgt gccgaatcgg tcgtccgcca tctccgcgcg 38100  
gcgggcgga gtcggatcgc gttccggacg ttcactgca ccggctcgcc gatgtcgccg 38160  
gcggcgattg catcccgat caaacgatgt ggcacggagt tgcgcatgtg gtggtttgtc 38220  
gcgaacacta cgcctgcctt ctccgcggca tcgacaagtt cttgtgcctc ggcgacggtc 38280  
agggcaagcg gtttttcgca caggacgtgc ttgccggctg cgatcgccgc aagtgcctga 38340  
tcgtgatgtg ccgagttgat actcgagatg tacacggcgt cgacgtcagg atcgttaagc 38400  
gccacggtga gttcgtcatg gcccggggcg ccgaatggtg ccgctacatc ctgcgctcgg 38460  
gccaggtcgc cgtgacgac gctgaccacc gcggcgtcag gctgggcgga gatcgctggg 38520  
cctacaaatg cgcgggcgat gtcgctggcg cccacgagcg accagcgga gggacggttc 38580  
gtcatgattc gatgggttcc tttcgggaat accgcgcctg tgcgacattt ccgctggcg 38640  
acagacgtga ggccagcgcg ggtggccgcg aggattacac gaccatgaaa cgggtcgccg 38700  
cttgactgtt agcaagatcg gaggtttcgc ataccgcgc tgatctcgt tgcccacgag 38760  
ctgtcctggt agtgccaggt cgtcaggatt ttgtcagcgc tccttcccca cccggctgca 38820

cgtctcccga gccggccttc cggatcgaac agcccagacat cgccgtcggc tggttgatg 38880  
 ccgtagcgtc ggcgtagctg tttccaggga gtacgccggg gtttgcgctc ctgccaccgg 38940  
 acgacctctt tccagagata cgatcgtagg tagctgaagg tcgcggttcgg gcatccatac 39000  
 ctgaagtacg ccttcaggc agcatccggg tccggatagc tttcaatgtg gggaaaagct 39060  
 gggtcgcccg cgacccgaaa accgacggca ccgctcgcg gcagcgtgca ccctggaggc 39120  
 gaggtccgtg ctgctgatcg tgatcccgtt ggcgagtggg tagccctgga tgaccgtaga 39180  
 tcagcgcacg taggtgatga gtttgggctc gggatcgctg acgtcagtga ccccgaggag 39240  
 cggacatggc ttcgggttgc gctgatctgg agacaggatg tttgtcgggg ttacacgttc 39300  
 tggagcagtc cgccatcgac accgacaacg gatccgggta tgtagctggc acgctcgctg 39360  
 gcgagaaatg tgaccgcgtc ggcgaattcc tcgcgccgcc cgtagcggcc gaccggaata 39420  
 gagctctctg attccctcgc gatgttctcc gcagtgcggg tctcccgtgc ggcctttgcc 39480  
 tgggtccaaact gggcgatgcg cgcagtggcg atgcgccccg gaacgacgat gttgctggtg 39540  
 atgcgctcgc ttgccacctc accggacagt gtcttgacc agcccaccag actcgagcgg 39600  
 agggcggttcg acaacccgag gttggggatg ggcgcatga cgccggacga agtcgacgtg 39660  
 atgatccggc cccatccccg ctcggtcatg cccgggagca cgcggtctgt gatcgcgacg 39720  
 acggagagga tcatcgctc gaagtgcga cgccacaggg tggcgctctg cccggacaca 39780  
 cgggtgggcg gcgggcccgc ggtgatgtt accaggatct ccaccgggccc gagctcgctc 39840  
 tcgacggcgg tgacgttctc atcgatgacc gacaggttcg ccaagtccca gaccaggctc 39900  
 atgttatcgc cggatcatgga cgctcgacc tccgtgagcg cgtccttgct gatgtcggcc 39960  
 agcgccacac gcgcgccttc ccggcttagc gaggcggcaa tggccgaacc gagtccgccg 40020  
 ccggctccga gtacgagggc gactcttcca ttgattgcga gatccatgac atgtccttgc 40080  
 ggtttcgggt gacggcggcc cggaaagtcg tggcacgccg ttgatggagg gacattggtg 40140  
 ggaggagtgg cagtggcccg gtcagccgcc taccacggcc tgagcgtga cctcgaacag 40200  
 gacaccgggt cggagaccgg cggatgatgt cgtccggggc gggccgggct cgctgaagta 40260  
 ctgcggtag atcgcgttca tcgctcgaa gtcgtccccg tcagcgagga acacattcac 40320  
 ggtgatcacg tctgcctcgc tgcagccggc tgaacgcagc gcggtcatga ggttctcgaa 40380  
 tgcccggcgc acctgaggtt ccagaccacc cgggtgcgagg gatctgtcgg cgaggtagcc 40440  
 gcaactgccg gcgattgcca cgatgttgcc gatgcgtacg gcggcggagt aagggcctcc 40500  
 cgcgggagga aggtcgacgt ctgtgaacgc ctgacgggac acggtgtcgg ggcctctctg 40560  
 tggctggtgg tcatgtccaa ctgacttcgg gagcttgtgg ggtcggttca accggagcct 40620  
 cccgtaggag gtctgagcc ccgtcagaag aacgtgagga cgccgtcgac gacgttgaag 40680  
 tcgtcgtcga tgatcgggat gaaggggagc ttgtcgaagg ccgtgcacgg gtgggacatg 40740  
 tcgaaagcga cgatgtctcc gggagcgtc gcgagcccgg gggcaacgcg aacgaaggcg 40800  
 tgctggtcca tcagcttgaa cgtctccgt tcccgcagcc gctgcctcga gccatccgcg 40860  
 cgaaacagct cgcgtggctg tggcagcgtc aggtcatggg ccgctcggc cttcccgatc 40920  
 gacaggatcg cgagatcagg ctccggccgg gacagcacgg aagcccatgc agtgagtgcg 40980  
 ttgtggagcc ggagcggttc ctgctcggcc cgtcggccag cgagcggcga gagctttctg 41040  
 tacttgccgg cgtcgtggga gatgtaacag ccgcttcgaa gcactaggct cactcgtcgt 41100  
 cggacaccgt cccacgagcc gagtcggcg acgacacgat cgaagtacga gcttccgcct 41160  
 gcggtcacga tgatgcggtt ggcgtcgaa aggtcgcgac gtgccagctc gagcacgatc 41220  
 gaccggacgc cagcgaagaa ggtgtcgagt gcatggaggt cctccggtgt catgcgcgcg 41280  
 gtcacgaggc cctcgtaggc ctgcagccg accagccgga gatgctcggg ccgcgcgacc 41340  
 gcccgggcga cctcgagcgc cgagtgcagg tctcgcacgc cacacctgcc accaggcaact 41400  
 ccgacctcga ggaggacgtt gagggtgacc ccgtggccgg agtcggcgat ggtatcggac 41460  
 atcgcccga ccgtgtccgc atcgtcgacc aggcagtaga agtcgaacga gctgtcgtg 41520  
 gcgatttcgg cagtcaccca acgcagcgc ggagcatcga cgagttcgtt ggccaggatg 41580  
 atgcgctgaa ccccgtagcg gcgcatcac gccacctgag tgggcgtcgc tgccgtgaac 41640  
 gccacgcac cgtgagacgc ctgaagggcg aagagctggg gtgacatggt ggtcttgccg 41700

tgcggggcgga	gaagcgcggt	gttgcgctcg	cagtacgcct	ccatagtcgc	caggttggcg	41760
ttcatagcgc	gtgcgtggag	tgtggtcacc	ggcagggcga	ggtctccgcg	ggccacgttc	41820
cagttctgct	ggccacctc	tccgaggagc	acacgagcac	ccaggggcac	gcccttgagt	41880
gcgtcgtcga	ggacgggtctc	gcggaggctcg	tccggcacatt	gcaggagtcg	ggtccggtcg	41940
agctgctcga	ccgcgtgagt	gtcatcgtg	tagtggcctt	tccgcgccgc	gtggtcagct	42000
gaccgatgcc	gtcgatcccg	cactccacac	ggtctccgc	gccgatgacg	agcgcgccag	42060
gggttccggt	ggagatgatg	tctcccgaa	acagcggcat	catcttcgag	tggaaagtca	42120
cgaggctggc	gggcttggtg	gtcatgtggg	cgaccgtgtt	ggtcctgaac	tcatcgccgt	42180
tcttgaccgt	cgagatctcg	atgtcgccca	agctttcgaa	cttctctatc	acctcatcga	42240
gcggtacgag	ctccggtccg	aagcagaaga	acgtccggaa	gttcttcgag	cggtgagat	42300
agcgcggatt	gatccgcaag	atgtcctcgg	cggtctggtc	caggatggtc	gtgaccccc	42360
acacgtagtc	catcgccctcg	gactcgctga	cgttctcgca	agtcctgccg	atgatgagcc	42420
cgatctcagc	ctcggacgtg	gtgcgtcac	tctgcgccgg	aatcacgac	ggctcatccg	42480
ggccgatgat	cgtgtggtcg	cacttcacga	aggacgccgg	ctggtcgggc	gccgactcgc	42540
tcaggctccg	tgcgtgatcg	ccgtagttca	gaccgatgcc	ccagatcttc	cggggggac	42600
ggtacggcgc	ggtgaaccgc	acctcgccca	cgcggtgtaa	ggcgaggtcg	ggggcggtct	42660
tgacacgctc	acgcagccgc	gcgaggtggt	ctcctgcgat	cacgtccatg	atggctcgtga	42720
ccgccagctc	gggggtctatc	tcccggattg	ctgcgacgcc	gcggtcgcctc	acgacaacgg	42780
ggatctctcc	ttccgcagcg	gtgatggctg	ccagggtgctc	ggtcacgaaa	gcccgctcct	42840
cggtcgaatg	gtgaggtagg	gggtggggga	ggtgggttaa	ctcgttccc	ggggcacgtt	42900
cccgcgctgc	gtgtgtgagg	agaggacaag	cgacatcacg	tggaggatgt	ggtcggcggt	42960
gttccgggct	gcccggctcg	cgtcgccagc	cagaacggcg	gatgccgtcg	cgatgtgctc	43020
cgaggcgcag	gagatggagt	caccggctga	gggagtgctg	cgttctcgga	gtacagcggt	43080
cagcaccagg	gcgttgctga	ccgcgcgctg	gaggatctcg	ttcccgcaga	actcggcgat	43140
gatgtggtgg	aagacgcgat	cctcggcgcg	gaactgggcg	gagtcgccgg	actctgccgc	43200
gcgcagggat	gaggtggcga	cctcgtggag	cctctcccgt	tgggcgctgg	tagcccgatg	43260
tgcggtgaac	caggctgccg	agctctccag	ccccgcgccg	tgctcgtagg	cgtcgcgaat	43320
ggcgcggacc	gaggcgctga	tgacgcgcag	ccccttcggg	gcagcctcga	ccaggcccac	43380
gtgccgcaag	cggagaagcg	cctcgcgcac	gggggtcttg	ctgacattca	gctcggccgc	43440
gagcttgccc	tcgtctcacac	gacgcccagg	tgggagtgcc	tgacttacga	tgctgtcgcg	43500
gatggcctcg	aagacccgat	ctgtcagatt	gtcgacgcgt	gcgtcgaggg	gtgcgagcat	43560
accgttgccg	gaccccaccg	cggagtctgc	ggcggctggg	cgatggtgca	gctccaccat	43620
gttcagcgtg	cccgcagtcg	gtggctgatt	cggttccatc	gtcccccttc	cttggaacagt	43680
ggaatacggg	atatggctct	caccgcgtaa	gcaaccgccc	gaggcgggac	tccgtagcgg	43740
atgaagtcct	tgggatatgt	cgtataccga	gggagagctg	atgaagcccg	agtcgcaccc	43800
ggccttcgag	agcctggatc	gactggccac	gatcgagatg	cgtccgccga	cgagcccggc	43860
cggcatcgtg	agcccgttct	acacgctggc	gcggggtgag	cggtcggagc	cgatcagcct	43920
tgcgatcgcg	cgctctctgc	tgcgtcgaca	gggtgaaaag	gcgttgatcg	tcacgggtct	43980
ggtcgaggaa	ggccggttcc	cgcaggggga	ggtggacggc	cccatcggct	ctctggcact	44040
cgcgcgcacg	ctcgccggte	ttggtagcga	ggtgaccatc	gtcatcgacc	cagaggccgt	44100
cggcccgggtg	gagaagctgg	tcgacgctgc	gggcctggat	ggcgtgacct	tgatggagtc	44160
ccgattctcc	agcgcggccg	aggcccggga	gttcgcggcg	cagttcggcg	tcgtcgccgc	44220
catcgagaag	ctggggcaga	actctgtttg	cggccgtcat	ctgatctggg	gcacgccggg	44280
caccgccggg	gatctgtttg	ccgacgacta	ccttcgcgcc	gccacggaca	acggtgctct	44340
gacgctggcc	gtgggagaca	acggcaacga	gatcgggttc	ggcaacattg	ccgccgagtc	44400
cgaggcgctg	actccgcgag	gcgtcagcgt	tgagggcgga	ttcttcgcgt	ccacgacggg	44460
ggatcatctg	ctgcctgcac	cgtgtcgaa	tttgggctgc	tacgtcatca	ccggtgcggg	44520
ggccatcctg	gctcgacgcg	caaacctggc	agtcaccggg	gatctggtgc	gcgagtggac	44580



cgagctgggg ctgcgcgcag gtctacgcag cggcggcggt gacgatcccg cgttccaagg 44640  
 cgacgacggg atccactgc gtttcgtcgc cgcgcgatgcc gagctgatct ccggaatcgt 44700  
 tcaccagagc ctgctcggcg acccctgggt ggcagaccag tgacggccac ccgtagcggg 44760  
 caaccctcgg ttgccagcct tcgtcatcat cgccgggaag gggaagtcac cgatggagtt 44820  
 cgttgatctg cagcacgcgg agcccatcga ggcattggcgt gccaaccccg cgaatgcgtg 44880  
 ggcgaggccc acgtccggga tctgtcatgg attcgtgcag gccaacctcg tgatgctgcc 44940  
 ccaggatgac gcgttcgact tcatgcgggt cgcccagcgg aaccccaagc cgtgcccgat 45000  
 cctcgaggtc accgacgcgg gagatccgga gccgaagctg accgcccctg gtgctgacct 45060  
 ccgcaccgac ctgccccaaat acagcgtgta ccgggaaggg gcgctgggtg aggagcgcag 45120  
 tgacgtcgtt gacctgtgga ccgacgacat ggtcgcgttc ctgcttggtc gcagcttctc 45180  
 cttcgagcac ttgctgctcg atgccggact ccgatccgg cacatcgacc aggggatcaa 45240  
 cggcccgggt tacctcagca accgtgagtg tgtgcccgca ggccagttcc gggggaagct 45300  
 agtggtttca atgcgtccca tgcccccgga ccgggtggcg gaggtgtca ccatacgcga 45360  
 tgcgcattccc gaggtgcacg ggggccccgt gcacatcggc tcccctgaga agttgggcat 45420  
 cgacaacggt gatgaacctg actggggtga acgagtgcgt atggagccgg gcgacgtgcc 45480  
 ggtgttctgg gcctgcggcg tgacccccga gcaggcggcg atcgaggcga agaccgagct 45540  
 aatgatcacg cactcgacgg gtcacatgtt cgtgaccgga atgcggatcg actccgagtg-45600  
 aggcgggggc cccgggggtg tgcc 45624

<210> 5

<211> 50000

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 5

gggcgcgcgc caattcgatg acgttcatgc gccgtgtcgg ggaatcgccg gtggcggcgc 60  
 cagcagaggc tgaacttact ggtgggtgtg ccaggaatcg gaggggcagt accgaatgag 120  
 cgaagccggg aacctgatag ccgtcatcgg actgtcctgc cgcctacccc aggcgcctga 180  
 cccggcttcc ttctggcggt tgctgcgcac cggaaacggac gccatcacca cggctccgga 240  
 agggcgggtg ggcgacccgt tgcttggtcg ggatgcgccc aaggggcccg aatgggggtg 300  
 ttctctggct gatgtcgact gcttcgatcc cgagttcttc gggatctcgc cgcgagaagc 360  
 ggcagccgtg gatccccagc agaggctggc tctggagctc gcctgggagg cactcgaaga 420  
 cgccggtatc ccgcgcggcg agctgcgcgg tactgcgcgc ggtgtgttca tgggggcgat 480  
 ctctgacgac taagccgccc tgctgcgcga gagcccgcgg gaagtggctg cgcagtagcg 540  
 cctcaccggc acccatcgaa gcctgatcgc caaccgcgtg tctatgtgc tcggcctgcg 600  
 cgggccaagc ctgacgggtg attcaggtca gtctcgtcc ctggtcggcg tgcatctcgc 660  
 cagcgagagc ctgcgacggg gtgagtgac gatcgcact gccggcggcg tgaacctcaa 720  
 cctggccgcc gagagcaaca gcgctctgat ggacttcggc gcgctctccc cggacggctg 780  
 ctgcttcacc ttcgatgtgc gggcgaacgg ttacgtccgt ggtgagggcg gcggccttgt 840  
 cgtgctgaag aaggccgac aggcgcacgc cgatggcgac cggatctact gcctcatccg 900  
 cggcagcgcg gtcaacaacg atgggggcgg tgccgggctc accgttcggg cggcggacgc 960  
 ccaggcggag ctgctgcgcc aggcataccg gaacgcgggc gtcgaccgg ccgccgtgca 1020  
 gtatgtcgag ctccacggca gcgcgaccag ggtcggggat ccgctcgaag cagcagccct 1080  
 cggagctgtc ctgggggcgg cgagacggcc cggcgacgag ctgcgtgtgg ggtcggcgaa 1140  
 gaccaacgtc ggccatctgg aagcagcggc gggcgtcacc gggttgctga agaccgact 1200  
 cagcatctgg caccgcgaac tgccgcccag tcttcatttc accgccccca acccggaat 1260  
 cccgctggac gaattgaacc tacgcgtcca gcgtgatctg cggccgtggc cggagagcga 1320



ggggccgctg	ctggccggcg	tcagcgcctt	cggaatggga	ggcacgaact	gccacctggt	1380
gctctccggc	acgtcccggg	tggagcgacg	gcgcagtggg	cccgtgagg	cgaccatgcc	1440
gtgggtcttg	tcggccagaa	caccggtcgc	attgctgctg	caggcggcgc	gcttgcacac	1500
gcacctcaat	acggccgggtc	aaagtccgtt	ggacgtcgcc	tactcaactg	cgaccactcg	1560
atccgcgctg	ccgcaccggg	ccgcgctggt	cgcggacgac	gaaccgaaac	tgtctgccgg	1620
gttgaaggcc	ctcgtgacg	gcgacgacgc	gcccacgctg	tgccacggcg	cgacttccgg	1680
cgagcgggca	gcgggtcttcg	tctttcccg	acagggcagc	cagtggatcg	ggatgggtag	1740
gcagctgctc	gaaacctccg	aggttttcgc	ggcgtcgatg	tcggactgcg	ccgacgcatt	1800
ggcgccgcac	ctggattggt	ccctgctgga	tgtgctgcgc	aacgcggccg	gcgctgcgca	1860
ccttgaccac	gacgatgtcg	tccagcccgc	gctgttcgcc	atcatggtct	cgctcgcgga	1920
gctctggcgt	tcgtggggcg	tgcgtccggt	ggcggctcgtc	gggcaactcg	agggggagat	1980
cgcggcggcc	tgcgtcgccg	gggccctgtc	cgtccgcgat	gccgccaggg	tgggtggcgt	2040
gcgcagcagg	cttctgacgg	cgctggccgg	cagtggcgcg	atggcctcgt	tgcagcatcc	2100
cgccgaagag	gtgcggcaaa	tcctgttgcc	ctggcgcgat	cggatcggcg	tggcgggggt	2160
gaacggaccg	tcgtcgaccc	tgggtgtcagg	ggaccgggag	gcgatggcgg	aactgctggc	2220
cgagtgcgca	gaccgagagc	tccgatgctg	ccggaattccc	gttgaatacg	cctcccatte	2280
gcctcacatc	gaggttgtcc	gggatgagct	gctggggctg	ttggcgccgg	tcgaaccag	2340
gacgggaagc	atcccgatct	attcgacgac	gaccggggac	ctgctggacc	ggccgatgga	2400
cgccgactac	tggtagccga	accttcgtca	accggtgctg	ttcgaagcgg	ccgtcgaggc	2460
cctgttgaag	cggggggtacg	acgcattcat	cgagatcagc	ccacaccggg	tgtgactgct	2520
gaacatccag	gaaaccgccg	tgcgagcagg	gcgggaggta	gtggcgctcg	ggacactccg	2580
ccgcggcgaa	ggtggcatgc	ggcaggcgct	gacgtcgtcg	gccagagcac	acgtacacgg	2640
agtggcccg	gactggcacg	cggtcttcgc	cggtaccggg	gcgcagcggg	tcgacctgcc	2700
gacgtacgcc	tttcagcgac	agcgctactg	gctggacgcg	aagcttcccg	acgtcgccat	2760
gcccagagagc	gacgtgtcga	cggcgttgcg	ggaaaagctg	cggtcttcgc	cgagggcgga	2820
cgtggactcg	acgacctca	cgatgatccg	ggcacaggca	gccgtggtcc	tcggccaactc	2880
cgatccgaaa	gaggtggacc	cggatcggac	gttcaaggac	ctgggcttcg	attcctcgat	2940
ggtggtcgag	ctgtgcgacc	gcctaaacgc	cgccacagg	ctgcgactcg	caccgagcgt	3000
cgttttcgac	tgtcctacgc	cggacaagct	cggccgccag	gtacggacgt	tgttgttggg	3060
cgagccggct	cccatgacgt	cacaccggcc	ggactccgat	gcggacgagc	ctatcgccgt	3120
gatcgggatg	ggctgtcggt	ttccgggtgg	ggtgtcctcg	cccgaaggat	tgtggcagtt	3180
ggtcgccgct	gggcgggacg	tcgtgtccga	gttcccggt	gaccgaggtt	gggacctgga	3240
gcgtgcgggg	acatcgcacg	tgcgcgccgg	cgggttcttg	catggcgccc	cggattttga	3300
ccccgggttc	ttccggattt	cgccgcgcga	ggcgttggcg	atggatccac	agcagcgggt	3360
gctgctggaa	atcgccctggg	aagcagtcga	acgaggcggg	atcaaccgcg	agcacctgca	3420
cggaaagtcaa	accggggtct	tcgtcggcgc	gacctccctg	gactacgggc	cacgcctgca	3480
cgaagcgctc	gaggaggcgg	ccgggtacgt	gctcaccggc	agcaccacga	gtgtggcgctc	3540
gggtcgggtt	gcgtattcgt	tcgggttcga	gggccctgcg	gtgacgggtg	atacggcggtg	3600
ttcgtcgtcg	ttggtggccc	tgcatttggc	gtgtcagtcg	ttgcgttcgg	gtgagtgtga	3660
tctggcgttg	gccggtggtg	tgaccgtgat	ggccacgcgc	gggatgttcg	tggagttttc	3720
gcggcagcgt	ggttttggcg	cggatgggcg	gtgcaagtcg	ttcgcggagg	ccgccgacgg	3780
caccggctgg	tccgagggtg	ctggccctggt	tctactggag	cggttgtcgg	atgcccgggc	3840
gaatgggcat	gaggtgctgg	cggttgttcg	tggtagtgcg	gtgaatcagg	acgggtgcgtc	3900
gaatggtttg	accgcgccga	atggttcgtc	gcagcagcgg	gtgattgccc	aggcattggc	3960
gagtgcgggg	ttgtcggtgt	ccgatgtgga	tgtgtgggag	gcgcatggga	cgggcacgcg	4020
gcttggtgat	ccgatcgagg	cgcaggcgct	gatcgccacc	tacggccagg	gccggcttcc	4080
ggaacggcca	ttgtggttgg	gctcgatgaa	gtcgaacatc	ggtcacgcgc	aggcagctgc	4140
ggggatagcc	ggcgtcatga	agatggtgat	ggcgatgcgg	cacgggcagc	taccgcgcac	4200

gttgcacgtg gatgagcggc cttctgggggt ggattgggtcg gcgggggggg ttcaactcct 4260  
 tacggagaac acgccctggc ccgggagtggt tcgtgttcgt cggtgggggg tgctgtcgtt 4320  
 cgggatcagt ggtactaacg cgcacgtcat cctcgaacag cccccgggag tgccgagtca 4380  
 gtctgcgggg ccgggttcgg gctctgtcgt ggatgttccg gtggtgccgt ggatggtgtc 4440  
 gggcaaaaca ccgaagcgc tatccgcgca ggcaacggcg ttgatgacct atctggacga 4500  
 gcgacctgat gtctcctcgc tggatgttgg gtactcgtcg gcgttgacac ggtcggcgct 4560  
 ggatgagcga gcggtggtgc tggggtcggg ccgtgaaacg ttgttgtgcg gtgtgaaagc 4620  
 gctgtctgcc ggtcatgagg cttctggggt ggtgaccgga tctgtggggg ctggggggccg 4680  
 catcgggttt gtgttttccg gtcagggtgg tcagtggctg gggatgggccc gggggcttta 4740  
 ccgggctttt ccggtgttcg ctgctgcctt tgacgaagct tgtgccgagc tggatgcgca 4800  
 tctgggccag gaaatcgggg ttccggaggt ggtgtccggt tcggatgcgc agttgctgga 4860  
 tcggacgttg tgggcgcagt cgggtttgtt ccggttgca ggtggccttg tgaagttgct 4920  
 ggattcgtgg ggggttcggc cgagtgtggt gttggggcat tcggtgggcg agttggcggc 4980  
 ggcgttcgcg gcgggtgtgg tgcgttgtc gggtgcggct cgggttggtg cggtcgtgc 5040  
 ccggttgatg caggcgttgc cgtctggcgg tgggatgctg gcggtgcctg ctggtgagga 5100  
 gctgttgtgg tcgttgttgg ccgatcaggg tgatcgtgtg gggatcgccg cggtaaacgc 5160  
 tgcggggctg gtggtgctct ctggtgatcg ggatgtgtc gatgacctt ccggtcggct 5220  
 ggacgggcaa gggatccggt cgaggtggtt gcgggtgtcg catgcgtttc attcgtatcg 5280  
 gatggatccg atgctggcgg agttcgccga attggcacga accgtggatt accggcgttg 5340  
 tgaagtgccg atcgtgtcga ccttgaccgg agacctcgat gacgctggca ggatgagcgg 5400  
 gcccgactac tgggtgcgtc aggtgcgaga gccgggtccg ttcgccgacg gtgtccaggc 5460  
 gctggtcgag cacgatgtgg ccaccgttgt cgagctcggc ccggacgggg cgttgtcggc 5520  
 gctgatccag gaatgtgtcg ccgcatccga tcacgcgggg cggctgagcg cggccccggc 5580  
 gatgcgcagg aaccaggacg aggcgcagaa ggtgatgacg gccctggcac acgtccacgt 5640  
 acgtggtggt gcggtggact ggcggtcgtt cttcgccggt acaagggcga agcaaactga 5700  
 gctgcccacc tacgccttcc aacgacagcg gtactggctg aacgcgctgc gtgaatcttc 5760  
 cgccggcgac atgggcaggc gtgtcgaagc gaagttcttg ggcgccgtcg agcacgaaga 5820  
 tgtggaatcg cttgcacgcg tattgggcat tgtggacgac ggcgctgctg tggattccct 5880  
 gagaagcgcc cttccggtgt tggccggttg gcagcgaacc cgcaccaccg agtccattat 5940  
 ggatcagcgg tgttacogaa ttggctggcg gcaggtagcc ggactcccgc cgatgggaac 6000  
 tgttttcggt acctggctgg tcttcgcgcc tcattggctg tccagcgaac cgaggtggt 6060  
 ggactgcgtt acggcactgc gggcacgttg tgcctcggtg gtgttggtgg aagctgatcc 6120  
 cgacccgacc tccttcggcg accgggtacg aacctgtgt tcgggccttc cggatcctgt 6180  
 tggcgtgttg tcaatgttgt gcttggaaga atcggtcctt ccgggatttt ctgcggtgtc 6240  
 acggggtttt gcgttgaccg tggagttggt gcgggttttg cgggcagctg gtgcgactgc 6300  
 ccggttgttg ttgctgacgt gtggtggcgt gtcggtggga gatgtaccgg ttcgtccagc 6360  
 gcaggccctg gcgtgggggt tggggcggtg tgtgggggtg gagcatccgg actggtgggg 6420  
 cggcttgatc gatattccgg tcttgttcga cgaagacgct caagagcggt tgcgattgt 6480  
 gctggcaggt ctcgatgagg acgaggtcgc gatccgtcct gacggcatgt tcgcgcgtcg 6540  
 gttggtacgc cacactgtct cagctgatgt gaagaaggcg tggcgcccca ggggatcgg 6600  
 gctggtgacg ggcggcacgg gtggtttggg ggcgcacgtt gctcgtggc tggccgacgc 6660  
 cggagccgaa catgtggcga tggtagtcg acgcggcgag caggcaccga gtgctgagaa 6720  
 gttgcggacg gaactggagg atctgggtac ccgggtgtcg atcgtgtcat gcgatgtgac 6780  
 cgatcgcgag gcgctcgccg aagtgtgaa agcccttccg gctgaaaacc cgttgaccgc 6840  
 ggtagtgcat gcggcaggcg tgatcgagac tggtagtcg gcggcaatga gcctggctga 6900  
 ttctgatcac gtgtgtccg caaaggtggc cgggtgcccg aatctggatg ccttgttggc 6960  
 cgatgtggaa ttggacgcgt tcgtcttgtt ctcatcggtg tcaggagttt ggggcgctgg 7020  
 gggacacggg gcttacgcag cggcgaatgc ctatctggat gcgctcgcg aacagcgtcg 7080

gtcgcgaggg	ctgggtcgca	ctgcggtggc	ctggggggccg	tgggcccggcg	agggcatggc	7140
ctccggagaa	acaggagacc	agctgcgccg	atacggcctt	tccccaatgg	ctccgcagca	7200
cgccatcgcc	ggaatccggc	aggccgtgga	acaggacgaa	atttccctgg	tagtgggcca	7260
tgtcgattgg	gcacgtttca	gcgcgggatt	gctggcggt	aggccgcggc	cgctgctgaa	7320
cgaactggcc	gaggtcaagg	aactcctcgt	cgatgccag	cccagggcgg	gagtccttgc	7380
cgacgcgtcg	ttggaatggc	ggcagcgatt	gtccgcggca	ccgaggccga	cacaggaaca	7440
gctgatactg	gagctggtac	gcggcgaaac	cgctctggtg	ctgggacacc	ccggggcagc	7500
ggccgttgca	tccgaacgag	ccttcaagga	cagcggattc	gactcgcagg	ccgcggtcga	7560
actccgcgtt	cggtcaatc	gagctaccgg	cctccagttg	ccatcgacaa	ttatcttcag	7620
ccatcccacg	cctgcggaac	tggctgcgga	gctgcgggcg	aggcttcttc	ccgagtccgc	7680
aggagcaggg	attcccagag	aggacgaggc	gcgaatcaga	gcggcactga	cgtcgatccc	7740
gttcccggcc	ttgcgcgagg	caggcttggg	gagtcgcgtg	ctcgcacttg	ccggacaccc	7800
ggtcgactcc	ggtatctcct	cggacgatgc	ggccgcgacc	tcgatcgatg	cgatggatgt	7860
agccggcctc	gtcgaagcag	cgctgggcca	acgcgagtcc	tgagaccgcc	gacctgggag	7920
atgacggtga	ccaccagtta	cgaagaagtt	gtcgaggcac	tgcgagcatc	gctcaaggag	7980
aacgaacgcc	tccggcgcg	cagggatcgg	ttctccgcgg	agaaggacga	tcccatcgcg	8040
atcgtggcga	tgagttgtcg	ttatcccggg	caggctctcct	cgccggaggga	cctgtggcaa	8100
ctgggtgcgg	gcggtgtgga	cgcatctcc	gaagttccgg	gggatcgcg	atgggacctg	8160
gatggcggtg	tcgttccgga	ctccgatcgt	cctggcacgt	cgatgcctg	cgccggcggt	8220
tttcttcagg	gcgtgtcgga	gttcgacgcg	ggtttcttcg	ggatttcgcc	gcgtgaggcg	8280
ctggcgatgg	atccgcagca	gcggttgcgt	ctggaagtcg	cgtagggagg	ccttcgagcg	8340
gctgggctgg	agcagcggtc	gacacgcggg	tcccgcggtg	gcgtgttcgt	cggcaccaat	8400
ggccaggact	acgcgtcgtg	gttcggaagc	ccgcgcctg	cggtggcagg	tcagtgtcgt	8460
acgggcgggtg	cggcagcggt	tcttccgggc	cggttgcgt	attcgttcgg	gttcgagggt	8520
cctgcggtga	cggtggatac	ggcgtgttcg	tcgtcgttgg	tggcgttgca	cctggcgggg	8580
caagcactgc	gggcccgtga	gtgcgacctt	gcccttgccg	gtggcgtcac	ggtgatgtcg	8640
acgcgaagg	tggtcctgga	gttctccgcg	caacggggtc	tcgcgcggga	tgggcgggtg	8700
aagtcgttcg	cggcgggtgc	ggatggcact	ggatgggggtg	aggggtgccg	actgttggtg	8760
ctggagcggt	tgtcggatgc	ccggcggaat	gggcatgagg	tgctggcggt	tggtcgtggt	8820
agtgcggtga	atcaggacgg	tgctcgaat	ggtttgaccg	cgccgaatgg	ttcgtcgcag	8880
cagcgggtga	ttaccagggc	gttggcgagt	gcgggggttg	cggtgtccga	tgtggatgct	8940
gtggaggcgc	atgggacggg	cacgcggctt	ggtgatccga	tcgaggcgca	ggcgctgac	9000
gccacctacg	gccgtgatcg	tgatcctggc	cgcccggttg	ggttgggggc	ggtcaagtcg	9060
aacatcggtc	atacgcaagc	ggcgccgggt	gtggctggtg	tgatcaagat	ggtgatggcg	9120
atgcggcacg	ggcagctgcc	acgcacgttg	cacgtggaat	cgccgtcgcc	ggaggtggat	9180
tggtcggcg	ggacggttca	actccttacg	gagaacacgc	cctggcccag	gagtggtcgt	9240
gttcgtcggg	tgggggtgtc	gtcgttcggg	atcagtggta	ctaacgcgca	cgatcatcct	9300
gaacagcccc	cgggagtgcc	gagtcagttc	gcggggcccg	gttcgggttc	tgctcgtggat	9360
gttccggtgg	tgccgtggat	ggtgtcgggc	aaaacacccg	aagcgctatc	cgcgcaggca	9420
acggcggtga	tgacctatct	ggacgagcga	cctgatgtct	cctcgtcgga	tggtgggtac	9480
tcgtcggcgt	tgacacggtc	ggcgtggat	gagcagcgcg	tggtgctggg	gtcggaccgt	9540
gaaacgttgt	tgtgcggtgt	gaaagcgctg	tctgccggtc	atgaggcttc	tgggttggtg	9600
accgatctcg	tgggggctgg	gggcccgcac	gggtttgtgt	tttccggtca	gggtgggtcag	9660
tggctgggga	tgggcccggg	gctttaccgg	gcttttcggg	tggtcgtgc	tgcccttgac	9720
gaagcttggtg	ccgagctgga	tgacatctcg	ggccaggaaa	tcgggggttcg	ggaggtgggtg	9780
tccggttcgg	atgcgcagtt	gctggatcgg	acgttggtgg	cgcagtcggg	tttggttcgcg	9840
ttgcaggtgg	gcttgctgaa	gttgctggat	tcgtgggggg	ttcggccgag	tgtggtggtg	9900
gggcattcgg	tgggcgagtt	ggcgccggcg	ttcgcggcg	gtgtggtgtc	gttgctcgggt	9960

gcggctcggg tgggtggcggg tgcgtgcccgg ttgatgcagg cgttgccgctc tggcgggtggg 10020  
 atgctggcgg tgccctgctgg tgaggagctg ttgtggctcg tgttgccga tcagggtgat 10080  
 cgtgtgggga tcgccgcggt caacgctgcg gggtcgggtg tgctctctgg tgatcgggat 10140  
 gtgctcgatg accttgccgg tcggctggac gggcaaggga tccggtcgag gtggttgccg 10200  
 gtgtcgcgcatg cgtttcattc gtatcggatg gatccgatgc tggcggagtt cgccgaattg 10260  
 gcacgaaccg tggattaccg gcgttgtgaa gtgccgatcg tgtcgacctt gaccggagac 10320  
 ctcgatgacg ctggcaggat gagcggggccc gactactggg tgcgtcaggt gcgagagccg 10380  
 gtccgcttcg ccgacgggtg ccaggcgctg gtgcgacacg atgtggccac tgttgctcgag 10440  
 ctcggtccgg acggggcggt gtccggcgctg atccaggaat gtgtcgccgc atccgatcac 10500  
 gccggggcggc tgagcgcggt cccggcgatg cgcaggaacc aggacgaggc gcagaagggtg 10560  
 atgacggccc tggcacacgt ccacgtacgt ggtggtgcgg tggactggcg gtcgttcttc 10620  
 gccggtacgg gagcgaaaca aatcgagctg cccacctacg ccttccaacg acagcggtac 10680  
 tggctggtgc catcggaattc cggatgatgt acagggtgcc gtctggccgg ggcggagcat 10740  
 ccgctggttg gtgctgtggt gccggtcgcg ggtggtgacg aggtggttgc gaccggcagg 10800  
 atttcggtgc ggacgcattc gtggtggcc gaacaccggg tgctgggtga agtgatcggt 10860  
 gcgggcaccg cgttgctgga gatcgccctg cagcggggg aacgtcttg tgtgaacgg 10920  
 gtggaagagc tcaccctgga agcaccgctg gtcctgcccg agcgcggggc gatccaggt -10980  
 cagctgcgag tgggcgcgcc cgagaattcc ggacgcaggc cgatggcgct gtattcacgc 11040  
 cccgaagggg cggcggagca tgaactggac cggcacgcca cgggcccgtt ggcgccaggc 11100  
 cgcggcgagg cggctggaga cctggccgac tggccggctc ctggcgcgct gccggtcgac 11160  
 ctcgacgaat tctatcggga cctgcgagag cttgggctgg agtacggccc gatcttccaa 11220  
 gggctcaagg cggcctggcg gcaaggggac gaggtgtacg ccgaagccgc gctgccggga 11280  
 acggaagatt ctggtttcgg ggtgcatccg gcaactgctg acgcggctct gcacgcaacg 11340  
 gctgtccgag acatggatga cgcacgcttg ccgttccagt gggaagggtg gtccctgcac 11400  
 gccaaaggcc cgccggcttt gcgggtccgc gtggtcccgg ctggtgacga tgccaagtcc 11460  
 ctgctggttt gtgatggcac cggtcgaccg gtgatctcgg tggaccgact cgtattgcgg 11520  
 tcggctgcgg cccggcgac cggtcgcgc cgacaggccc atcaagctcg gttgtaccgg 11580  
 ttgagctggc caacggttca actgccgaca tccgctcagc caccgtcctg cgtgcttctc 11640  
 ggcacctcag aagtgtccgc tgacatacag gtgtatccgg acctccggtc gttgacggct 11700  
 gcgttggatg ccggtgccga accaccggc gtcgtcatcg caccacgccc ccccgcggt 11760  
 ggacgaacag cggatgtccg ggagacgact cggcatgcac tcgacctggt acaaggctgg 11820  
 ctttccgac agcgactcaa cgaatccga ttgctcctgg tgacacaggg agcagtggcc 11880  
 gtggagccgg gcgaaccctg gaccgatctg gcgcaggccg cgctctgggg actgctgcgg 11940  
 tcgacgcaga ccgaacaccc tgatcgcttc gtcctcgctg atgtgcctga gcccgcgcaa 12000  
 ctccctcccc cgctgccggg ggtgctggcc tgcggcgaac ctgagctcgc gttgcgacgt 12060  
 ggccggcgctc atgcgccag actggctgga ctgggcagcg atgacgtcct gcccggtccg 12120  
 gacggcaccg ggtggcgatt ggaggccacg cggccgggaa gcctggatgg gttggcattg 12180  
 gtggacgaac cgacggccac ggcaccgctg ggtgacgggt aggtcaggat tgcgatgcgc 12240  
 gcggccgggg tgaacttccg ggtgcgctc atcgcgctcg gtatgtatcc cgggtgtggca 12300  
 tcgctgggca gtgagggcgc cggggtcggt gtggagaccg gccccggcgt caccggcctg 12360  
 gcaccggcg accgctgat gggaatgat ccgaaggcgt tggggccgct cgcggtcgcc 12420  
 gaccatcgca tggtgacgag gattcccgct ggttggagct tcgcgcgggc cgcacgggtg 12480  
 ccgatcgtct ttctaccgc ctactacgcg ctggttgatc tcgcggggt gagaccaggg 12540  
 gagtcgttgc tggttcattc ggccgcgggt ggggtgggga tggccgcgat ccaactcgcc 12600  
 aggcaacctg gtgcagaggt gtacgccacc gctagcgagg acaagtggca agccgtggag 12660  
 ctgagccgag aacacctcgc ttctgcgcg acgtgcgatt tcgagcagca gttcctcggg 12720  
 gcaaccggcg gacgcggcgt cgacgtcgt ctcaactccc tcgccgggga gttcgccgat 12780  
 gcgtctctgc gaatgctgcc gcgcgggtggc cgtttcctgg agttggggaa gacggatgtt 12840

cgtgaccccg	tcgaggtcgc	cgatgcgcac	ccgggcgtgt	cttaccaggc	tttcgatacc	12900
gtagaggcag	gcccgcagcg	aatcggcgag	atgcttcacg	agctggtgga	gttggttcgag	12960
ggacgcgtgc	tggagcccct	gcctgtcacg	gcttgggacg	ttcggcaggc	gcccgcagcg	13020
ctacggcacc	tgagccaagc	gcggcatgtg	ggaaagctgg	tgctcaccat	gcctccggtg	13080
tgggacgccg	caggcacggt	tctggttacc	ggcggaacgg	gagcaacttg	cgcagaggtc	13140
gcccgcaccc	tcgtgatcga	gcgcggggtg	cgaaacctgg	tcctcgtcag	caggcgcggt	13200
cccgcagcca	gtggcgctgc	tgagctcgtg	gcgcaactga	cggcctacgg	tgccgaggtt	13260
tccttgacag	cttgcgatgt	cgccgatcgt	gagaccttgg	cgaaggtgct	tgccagcatc	13320
ccggacgagc	atccgttgac	cgcctggtg	cacgcggctg	gtgttctcga	cgcagaggtg	13380
tccgaatcgc	tcaccgtgga	gcggctggac	caggttctgc	gcccgaaggt	cgatggcgcg	13440
cggaatctgc	tcgagctgat	cgacccggac	gtggccctcg	tggtgttctc	gtcgggtgctg	13500
ggtgtgctcg	gcagcggtgg	gcagggtaac	tacgcggcgg	ccaactcctt	cctcgacgca	13560
ttggcgagc	aaaggcagtc	gcgcggccta	ccgacgagat	cattggcctg	ggggccctgg	13620
gcggaacatg	gcatggccag	caccttgcg	gaagccgagc	aggatcgatt	ggcgcgatct	13680
gggttgctgc	cgatctcgac	cgaggagggg	ttgtccaggt	tcgacgccgc	gtgcggcggc	13740
gcgcataccg	tggtggcgcc	ggttcgattc	agccgcttgt	ccgacgggaa	cgcgatcaag	13800
ttctccgtcc	tgcaaggttt	ggtcggggcg	catcgcgctc	acaaagcggc	gactgcggat	13860
gatgccgaga	gcctccggaa	acggttgga	cgcttgccgg	atgcagaaca	acatcggatt	13920
ctgctggacc	tcgtccgcat	gcatgtggcg	gcagtgtctg	gattcgccgg	ttctcaggag	13980
atcaccgcgg	acggcacggt	caaggtgctg	ggcttcgact	cgttgaccgt	ggtcgagttg	14040
cgcaaccgga	tcaacggggc	gacggggctg	cgactgcccg	ccaccctggt	gttcaactac	14100
ccgacgccgg	atgcgctcgc	cgcgcacctc	gtcaccgcgc	tgtccgcaga	ccgcctggcc	14160
gggacattcg	aggaactcga	caggtggggc	gcgaacctgc	ccacgctggc	cagggatgag	14220
gccacgcggg	cgcagatcac	caccgggcta	caggcgatct	tgcagagcct	ggcggacgtg	14280
tccggcgga	ccggcgggcg	ctccgtgccg	gaccggctca	gatcggccac	ggacgacgag	14340
cttttccaac	tcctcgacaa	cgatctcgaa	cttccctgat	gcctcagccg	gagccttcgc	14400
aacttctctg	agggaaacgc	cacatgtcga	atgaagagaa	gctccgggag	tacttgccgc	14460
gtgcgctcgt	ggatctgcac	caggcgcgcg	agcggctgca	cgaggcggag	tcgggagagc	14520
gggaacccat	cgcgatcgtg	gcgatgggct	gccgggtacc	gggtgggggtg	caggaccggg	14580
aagggtctgt	gaaactgggt	gcctccgggt	gcgacgccat	cggtgaattc	cccgtgatc	14640
gtggttgga	cctcgacgag	ctctacgac	ccgacccgga	tcagcccgga	acctgctaca	14700
cccggcacgg	cggcttcctc	cacgacgccg	gcgagttcga	cgcgggattc	ttcgacatca	14760
gcccccgta	ggcgctcgcg	atggaccgcg	agcagcggt	gctgctggaa	atctcctggg	14820
agaccgtcga	atccgctggg	atggaccgga	ggtccttgcg	ggggagccgc	accggggtgt	14880
tcgcgggatt	gatgtacgag	ggctatgaca	ccggcgccca	ccgggcagga	gaaggtgtcg	14940
aaggctatct	cggaaccggc	aatgcgggaa	gcgtcgccct	tggtcgggtt	gcgtatgcgt	15000
tcgggttcga	gggcccagcg	gtgacggtag	acacggcggt	ctcgtcgtcg	ttggtggcg	15060
tgcatattgg	gtgtcagtcg	ttgcggcagg	gcgagtgtga	tctggcgctg	gccggtggag	15120
tgacggtgat	gtcgacgccg	gagaggttcg	tggagttctc	ccgtcagcgt	ggtctcgcac	15180
cggatgggcg	gtgtaagtcg	ttcgcgggcg	ctgcggatgg	aaccggttgg	ggtgaggggtg	15240
ccggtttgg	gttgctggag	cggctgtcag	acgccaggcg	gaacgggcat	cgggtactgg	15300
cggttgttcg	tggtagcgcg	gtgaatcagg	acggtgcgtc	gaacggattg	acggccccga	15360
acgggctggc	ccaggagcg	gtcattcagc	aggtgtcac	gagtgcgggg	ctgtcggcgt	15420
ccgatgtgga	cgtgtggag	gcgcatggaa	cgggtacgcg	gcttggtgat	ccgatcgagg	15480
cgcaggtct	gatagccgcc	tatggacagg	atcgggaccg	ggaccggccg	ctgtggttgg	15540
ggtcggta	gtccaacatc	ggtcatacgc	aggcggtgc	gggcgtcgt	ggtgtgatca	15600
agatggtcat	ggcgatgcgg	cacggggagc	tgccgcgcac	gttgacagtg	gacgagccga	15660
attcgcacgt	ggactggtcg	gctggtgcgg	tccgactcct	gaccgagaac	atccgctggc	15720

cagggacggg tacgcgccgc gctggagtggt cgtcgttcgg ggtaagcggg accaacgcac 15780  
acgtcatcct cgaacacgac ccgctcgccg tgaccgagaa cgaggaagca gcgcagtcct 15840  
cagcacctgg gatcgtgcc tgggcgttgt ccgggcggtc gtcgacggcg ctgcggggcc 15900  
aggccgaacg gctgcgcgag ctgtgcgagc agaccgatcc cgaccccgct gatgtcgggt 15960  
tctcactggc cgccacgcgc acggcttggg agcaccgagc ggtggtgctt ggtcgggaca 16020  
gcgctacgtt gcgctccggg cttggcggtt ttgccagcgg tgaaccagcg gtcgatgtcg 16080  
ttgaggggag cgtcctggac ggcgaggtcg tcttcgtctt ccccggtcag ggctggcagt 16140  
gggcccgtat ggcagtcgac ctgctggacg cttcgccgac gttcgcgcg caccatggacg 16200  
agtgcgccac cgcgctgcgg aggtacgtgg actggtcgtt ggtcgacgtg ctgcgcggag 16260  
cggagaactc cccaccgctg gaccgggtgg acgtgctcca gcccgctcc ttcgcggtga 16320  
tggtgtcgtc cgccgaggtg tggcgcttct acggggtgag gccggcgggc gtcgtcggcc 16380  
acagtcaagg cgaaatcgcc gcggcctgcg cagccggggg gctgccgctg gaggatgcgg 16440  
ccaggcttgt cgcattgcgc agcagagcgt tgaagggaact ttcggggcgg ggtggcatgg 16500  
cgtcgtcggc ctgccctgcg gatgaggtcg cggcattggt cgcgggatcg ggcggccgtc 16560  
tggaagttgc ggcgatcaac ggcccgcgat cggtcgtggt gtccggcgat ctggaagcgg 16620  
tggaagaaact gctggcagag tgcgctgaaa aggacatgcg tgcacgcgct atccccgtcg 16680  
actacgcctc gcattcagcg cacgtggagg tgggttcggag cccgggtgctg gcggccggccg 16740  
ccggggtgcg acaccgggac ggccaggtgc cgtggtggtc gacggtgatc ggcgactggg 16800  
tggtatccggc caggctggac ggcgagtatt ggtatcgaa cctccggcag ccgggtccgg 16860  
tcgaacacgc cgtgcagggc ctggtcgagc ggggattcgg cctgttcacg gaaatgagt 16920  
cgcacccggt gctgaccacg gcggtcgagg aaaccgggtc ggagtcggag accgccgtgg 16980  
ccgcggtagg taccttgcca cgtgactcgg gcggcctccg gaggttggtt cattcgtgg 17040  
ccgaggcgta cgtgcgcggc gccaccgtgg actgggcccgt ggcgttcggg ggcgcggggc 17100  
gacggctgga cctgccgacc taccggttcc agcgcacagc gtactggctg gacaaggag 17160  
ctgcctccga cgaggctcgt gcggtctcgg acccgcgggc gggctggttc tggcaagccg 17220  
tggcgcgcca agacctgaaa agcgtgtccg atgcccctga tctcgacgcc gacgcaccgc 17280  
tgagcgcaac acttcacgac ctgtccgtct ggcaccgtca ggaacgagaa agggctcttg 17340  
cagacgggtg gcggtaccga gtcgactggg tacgggtggc cccgcagccg gtccggagaa 17400  
cgcgggaaac ctggtcctcg gtcgttcccc cgggcggcat cgaggaagcg ctggtcgaac 17460  
ggctgacgga tgcgttgaa acgcgagga tcagaccct gcgcctcgac gtgccaccgg 17520  
cggcgaccag tggcgaactc gcaaccgaac tccgcgcgc agccgacgtt gaccgggtga 17580  
aggcaatcct gtcgtcacc gcgttgagc agcgaccca ccccgaaatgc aaggacgtcc 17640  
cgagcgggat tgccttgctg ctgaacctgg tcaaggcgt cgggtgaagc gacctagaa 17700  
ttcctctgtg gaccatcacg cgtggtgcgg tcaaggcagg cccgcagat cggctgctgc 17760  
gcccgatgca ggcgcaagca tggggtctgg ggcgagtagc cgcactcgaa cccccgagc 17820  
gctgggggtg gctgatcgac ctgccggatt cgtgagcgg cgacgtcctc acgaggtctg 17880  
gcgaagcgt caccaacggc ttggcggaag accaactggc gattcgccag tcgggcgtgc 17940  
tggcccggcg actggtaccc gcccgcgga atcagcccgc tggacgtaag tggcggcccc 18000  
gagggagcgc gctgatcacg ggcggactcg gcgcggtgg cgcacaggtg gcgaggtgg 18060  
tggccgaaat cggagccgag cgaatcgtgc tcaccagtc agggggcaac caagcagcag 18120  
gcgcgcgga gctggaagcc gaactccgg cccttgagc gcaagtgtcc atcgtggctt 18180  
gcgacgtgac cgatcgtgac gagatgtcc cactactggc cgagttcgac gtcaccgcg 18240  
tggtccacgc ggccggagtc ggtcggctgc tgccgttggc ggagaccgac cagaacggcc 18300  
tggccgaaat atgcgcggcg aaggtccgc gcgctcaggt gctggacgaa ctgtgcgaca 18360  
gcaccgatct cgatgccttc gtcctgttct cctcgggtgc cggggtatgg ggcggggcg 18420  
gtcagggcgc ttacggcgcg gcgaacgc tcttgacac actcgccgaa caacgccgag 18480  
cacgcggtct gccggcaacc tcgatctct ggggcagttg ggcggcggc ggcagggccg 18540  
acggcgcggc gggcgaacac ctgcggcgac gcgggatacg tccgatgccg gcggcgctcg 18600

ccatcctggc tctgcaggaa gtacttgacc aggatgagac gtgcgtgtcg atcgctgatg 18660  
tggactggga ccgattcgtt cccacgttcg ccgcgactcg cggcaccggt ttgttcgacg 18720  
aagtgcgggc ggcgagaaaag gcgatgcccg cgaatgggccc ggcagaacca ggcggctcgc 18780  
cgttcgcccg caatctcgcg gagctgccgg aagcccaacg acgccacgaa ctggtggatc 18840  
tggtgtgcgc ccagggtggca accgtgctcg ggacggcgag tcgcgaggaa gtccagcccg 18900  
agcgggctgt ccgcgcgctc gggttcgact ccctcatggc ggtggatctg cgcaatcgtt 18960  
tgaccaccgc caccgggttg cgctgcccga ccacaaccgt cttegactac ccgaatccgg 19020  
ccgccttggc cgctcacctg ctcgaggagc tgggtgggtga tgtcgcgctg gctgcgggtga 19080  
ccgctgccag cgcgcccgcg agtgacgaac cgatcgcgat cgtecgcatg agctgccggt 19140  
ttccgggttg cgcgcactcg ccggaagacc tgtggcggtt ggtcgccgccc ggcacggagg 19200  
tgatcggcga gttcccctcc gaccgggggt gggatgcgga aggcctttac gatccggatg 19260  
cttcaggcc tggaacgacg tatgcgcgga tggcgggatt cctctacgac gccggtgagt 19320  
tcgatgccga cctgttcggc atcagcccac gtgaggcgtt ggcgatggat ccgcagcagc 19380  
ggttggtgct cgaaatcgcc tgggaagccc tcgaacgggc cggaatcgat ccgttgtcct 19440  
tgaaggcag tgggtcggc acgtacatcg gcgctggaag ccgtgggtac gcgacggatg 19500  
tgccgagtt tcccgaggag gcgagggtt acctgctgac gggtagctcg gccagtgtgc 19560  
tgtcggtcg ggtcgctat tcgtttggtt tcgagggtcc tgcggtgacg gtggatacgg 19620  
cttgttcgtc gtggttggtg gcgttgcatc tggcgtgcca gtcgttgctg tcgggcgagt 19680  
gtgatctggc gttggccggt ggtgtgaccg tgatgtcgac gccggagatg ttcgtggagt 19740  
tctcccgta gcgcggtttg gcgcgggatg ggcggtgcaa gtcgttcgcg gagagcgcg 19800  
acggcaccgg ctggggcgaa ggcgcgggccc tgttgttgct ggagcgggtg tcggacgccc 19860  
accggaatgg gcatcggtg ttggcggtgg ttcgtgggtc agcgggtgaat caggacggcg 19920  
cctcgaacgg actggcggcg ccgaacggtc cgtecgagca gcgggtgatc aaccaggcac 19980  
tcgcgaatgc ggctctttcg gcgtccgatg tggatgcggt ggaggacat ggcaccggga 20040  
ccaggtcggg tgatccgatc gaggcgcagg cattgatcgc aacgtatggg caggcccggg 20100  
agcgggatcg gcccttggtg ctggggctcg tcaagtcgaa catcggtcat acgcaggccc 20160  
cggcgggtgt tgccggtgtg atcaagatgg tgatggccat gcggcacggg cagctgcccg 20220  
cctcgctgca cgcggatgag cccacgtcgg aggtcgattg gtcgtcgggg gcggtccggc 20280  
tcctcgccga acaggtacct tggccggagt ctgaccgtgt tcgtcgggtg ggggtttcgt 20340  
cgttcgggat cagcggcacc aacgcacatg tgatcctcga acaagctacg aatgcgccag 20400  
atagtacagc ggagacggac aaaacagaat ccgatctac tgtcgatatt ccggtcgttc 20460  
cctggttggt gtcgggaaag acgacggatt ccctgcgggg acaagccgaa cgagtcttgt 20520  
ctcaggtcga gtcccggccg gagcagcgtt cgctggatgt tgcctactcg cttgcttctg 20580  
gccgagccgc gctggatgaa cgcgctgtcg tgcgtgggtg ggaccgcggt gagctggttg 20640  
ctggactggc ggcgttggcc gccggtcagg aggtctctgg ggtgatcagc ggaactcgtg 20700  
cttctgctcg gttcgggttc gtgttctcgg ggcagggttg tcagtgggtg gggatgggca 20760  
gagcgtctta ctcgaagttt ccggtgttcg ctgctgcgtt tgatgaggct tgcgccgagt 20820  
tgagggcaca tctgggggaa gaccgcccgg ttcgggatgt ggtcttcggt tccgatgcgc 20880  
agctgctgga tcagacgctg tgggcgcagt cgggtctgtt cgcgctgcaa gccggcctct 20940  
tggggctgct gggttcgtgg ggcgttcggc cggatgtggt gatggggcat tcggtcgggg 21000  
agttggccgc cgcgttttgc gctggcgtgt tgcgttgctg ggatgcggct cggttggttg 21060  
ccgcgcgcgc ccggttgatg caagccctgc cctctgacgg cgcgatgttg gcggtggctg 21120  
ctggtgaaga ccttgttcgg ccattgctgg ccggtcggga ggagtccgtg agcgtcgccg 21180  
cgctcaatgc ccccggttcg gtggtgttgt cgggcgatcg ggaggtgctg gccagcatcg 21240  
tcggccggct gaccgagctc cgagtccgga cgcggcgtt gcgggtctcc catgcttttc 21300  
attcgaccg gatggaccg atgttgggcg agttcgcca gatcgccgag tctgcggagt 21360  
tcggtaagcc aacgacacc cttgtgtcga cgttgacggg tgagctcgac agagccgcg 21420  
aaatgagcac accagggtat tgggtgcgcc aggcgcgtga acccgctcgt ttcgccgacg 21480



gtgtccaggc cctggcagcg cagggcatag gcacggtcgt cgagctcggc ccggacggaa 21540  
 cgctggcggc actgggttcgg gagtgtgcga ccgagtcgga tcgggttggg cggtatttcgt 21600  
 cgatccact gatgcgagc gagcgggacg agaccggttc ggtgatgaca gccctggcgc 21660  
 atctccacac ccgtgggtggg gaggtggact ggcaggcgtt tttcgcgggt accggcgcta 21720  
 ggcagctcga gttgccaacg tatgccttcc aacgacagca ctactggatc gagtccagtg 21780  
 cgcgggccagc acgcgaccgc gcagacatcg gcgaggtggc ggaacagttc tggaccgcgg 21840  
 ttgaccaagg cgatctggca acgttgggtcg ccgctctgga tcttggggcg gacgacgaca 21900  
 catgcgcatc gttgagcgat gtattgccgg cgttgctctc ctggcgaagc ggactccgca 21960  
 accgttcgct cgtcgattcc tgccgggtacc gaatcagttg gcattcctct cgggaggtgc 22020  
 cggccccgaa gatttcgggt acctggctgt tggctgtgcc cgggtgctgcg gatgacggat 22080  
 tggtcacggc tttgacgagt tcaactggctg gaggcggcgc cgaggtcgtc cggatcggcc 22140  
 tgtccgaaga ggacccgcac cgcgaggacg tcgcacagcg gctggccaat gcgctgacgg 22200  
 atgccggtca actcgggtggc gtgctttcgc tgttggggct cgatgaatcg cctgctccgg 22260  
 gattctcctg cttgccaaact ggtttcgcgc tgactgtgca gcttctgcgg gccttgccga 22320  
 aggcgcagct cgaggcgctt ttttggggcg tgacgcgcgg cggcgctcgcg ttggaagatg 22380  
 tacgctgtc tccggagcag gccctggtct gggggctgct gcgtgtcgcg ggactggagc 22440  
 acccgaggt ctgggggtggc ttgatcgacc tgccatcgga ctgggacgac cgattgggtg 22500  
 cccggttggc ggggtgtgtg gcggatggtg gcgaggatca agtcgccatt cgcctgggtg 22560  
 gtgtgttcgt gcggcggttg gaacgcgctg gtgcgtcggg tgccgggtcg gtgtggcgtc 22620  
 ctccgggggac ggtgttgggt acgggtggta cgggcgggtt gggggcgcat gttgcccggt 22680  
 ggttgggcgg tgccggggct gagcacgtgg tgttgaccag ccgtcgagga gcggacgctc 22740  
 cgggcgctgg ggaattgcgg gcggagctgg aggcgctggg tgctcgggtg tcgattgtgc 22800  
 cctgcgacgt ggtgatcgt gacgcagtgg ctggagtgtt ggcagggatc ggtggggagt 22860  
 gtccgctgac tgcggtggta cacgccgccg gggctcggcg ggcgggcgac gtagtgaga 22920  
 tgggtttggc ggattttgca gcggtgtgtt cggcgaaggt gcgtggtgcg gcgaatctgg 22980  
 acgagttgct ggccgactcg gagctggatg cgtttgtgat gttctcctcg gtgtcggggg 23040  
 tgtggggagc cggcgacag ggtgcgtatg cggctgcgaa cgcctacttg gatgcgttg 23100  
 ccgagcagcg tcgggcgagg ggattggtcg ggaccgcggg tgctggggga ccgtgggccc 23160  
 gtgacggcat ggccgccggc gaaaccggcg cacagctgca ccgatgggc ctggcgtcga 23220  
 tggaaaccgag cgcggcgctg ctggcacttc aggtgcat ggaccgcgat gagacctccc 23280  
 tcgtcgtggc cgatgtcgat tgggcacggg tcgccccagc cttcacctcg gcacgtcgac 23340  
 gcccgctgct ggacaccatc gacgaggccc gagccgcatt ggaaaccacc ggcaacaag 23400  
 cgggcacagg caaaccggtt gagctgacgc aacgcctggc cggactgtcg cggaaggaa 23460  
 gcgacgatgc ggtattggat ctggtgcggg cggagacggc ggtgtgctg ggacgcgacg 23520  
 atgccacggc cctggcgcca tcgcgccgt tccaggaaact cggattcgac tccttgatgg 23580  
 cgggtggagct gcgcaaccgg ctgaacaccg ccaccgggat ccagctgccc gccagcacga 23640  
 ttttcgacta cccaatgcc gagtcgctgt cgcgtcacct ctgcgccgag cttttcccaa 23700  
 cggagactac cgtggactcg gcccttgccg agctcgatcg aatcgagcag cagctctcga 23760  
 tgctcaccgg cgaagcgcg gcacgggacc gaatcgcgac acgactgca gccctccacg 23820  
 agaagtggaa cagcgagct gaagtaccga ccggagccga tgtcctgagc acgctcgatt 23880  
 cggcgacgca cgacgagata ttcgagttca tcgacaacga gctcgacctg tcttgagcag 23940  
 ttcttgccga acttcaagcg ccgaaatcgg gtggaaatca caatggccaa tgaagaaaag 24000  
 ctcttcggct atctgaagaa ggttaactgcg gacctgcac agaccggca gcgcctgctc 24060  
 gcggccgaga gccggagtca ggagccgatc gcgatcgtct cggcgagctg ccgactgccc 24120  
 ggcggcgctg actctcccga agcgtcttg caactcgtgc gactggcac cgacgccatc 24180  
 tcggagttcc ccgccgaccg gggctgggat ctcggccggg tgtacgatcc cgaccgaac 24240  
 caccaggga cgctgtacac gcgggcccgg ggtttcctcg caggagcggg cgatttcgac 24300  
 cccgccatgt tcgggatttc gccgcgtgag gcgttggcga tggaccgcga gcaacgggtg 24360



ttgctggagc	tgctcctggga	ggccctcgaa	cgggcgggca	tagacccgac	atccctgcgc	24420
ggcagcaaga	cgggtgtctt	cgggtggtgc	acgccccagg	agtacgggcc	gtccttgacg	24480
gagatgagcc	gaaacgctgg	gggtttttgga	ctcaccgggc	ggatggtgag	tgtggcgctc	24540
ggtcgggttg	cgtattcgtt	tggtttttgag	ggtcctgcgg	tgacggtgga	tacggcggtg	24600
tcgtcgtcgt	tggtagccct	gcatttggcg	tgtcagtcgt	tgcgttcgcg	cgaatgcgat	24660
ctcgcgctgg	ccggcggtgt	gacggtgatg	gcgacaccgg	cgacgttcgt	ggagtctctc	24720
cgtcagcgtg	gtttggctcc	ggacgggcgg	tgcaagtcgt	tcgcggtcgc	cgcggatggc	24780
accgggtggg	gtgaggggtg	cggctctggt	ttgctggagc	ggttgctcga	tgcgcggcgg	24840
aatgggcacg	aggttctggc	ggtggtgcgg	ggtagcgcg	tgaaccagga	cggcgcgctc	24900
aatggtttga	ctgcgcgcaa	tggtagctcg	cagcagcggg	tgatcaccca	ggcgttggcg	24960
agtgcggggc	tgctcggtttc	cgatgtggat	gcggtcagag	cacatggggac	cgggaccacg	25020
ttgggtgatc	cgatcgaggc	acaggccctg	atcgccacgt	acgggcaggg	ccgggagaa	25080
gatcgggcgt	tgtggttggg	gtcgggtcaag	tccaacatcg	gtcacacgca	ggcgcccgct	25140
ggcgttgccg	gcgtcatcaa	gatggtcttg	gcgatgcggc	acgggcagct	gcccgccacg	25200
ttgcatgtgg	atgagcccac	gtcggcggtg	gactggtcgg	cgggttcggt	ccggcttctc	25260
acggagaaca	cgccctggcc	ggacagtggg	cgtccttgcc	gggtgggggt	gtcgtcgttc	25320
gggatcagcg	gcaccaacgc	acatgtgatt	ctcgaacagt	ctccagtcga	gcagggcgaa	25380
ccggccgggc	cggctgaagg	cgagcgggaa	ccggatgtag	ccgtccccgt	ggtgccttgg	25440
gtgctgtcgg	gtaagacacc	ggaggctgcg	cgggcgcagg	ccgaacgggt	gcattcgcac	25500
atcgaggacc	ggccggggct	gtcgcgggtg	gatgtggcgt	attcgctagg	aatgacacgc	25560
gcggcgctgg	atgaacgcgc	agtgggtgtg	ggctcggacc	gtgccgcgct	cctgaccggg	25620
ttgagggcat	tcgccgacgg	ctgcgatgcg	cccgaagtgg	tttcggggtc	tgtggggctt	25680
ggtggcccg	tcgggttcgt	gttctcgggt	cagggtggtc	agtggccggg	gatgggccgg	25740
gggtctact	cgggtgtttc	ggtgttcgcc	gacgcgttcg	acgaggcttg	cgcggagtgt	25800
gatgcacacc	tgggccagga	actgcgggtt	cgggatgtgg	tgttcgggtc	gcaagcgtgg	25860
ttgctggatc	ggacggtgtg	ggcgacgtcg	ggtttggttc	cgttgacgat	tggcttgctg	25920
cggctgctgg	gttcgtgggg	tgttcggccg	gatgtggtgt	tggggcactc	ggtgggtgag	25980
ctggctgcgg	tgcatgcggc	tgggtgtgtg	tcgttgctcg	aggccgcgcg	gttgggtggc	26040
ggtcgcgccc	ggttgatgca	ggcgttgcc	tctgggtggtg	ccatgctcgc	ggtcgtacg	26100
ggtgagtttc	aggtagatcc	tctgctggat	ggggtgcggg	accggatcgg	tatcgcgcg	26160
gtgaatggcc	cggaatcggg	tgtgctctct	ggtgaccgcg	agctgctcac	cgagatcgt	26220
gatcggttgc	acgatcaggg	gtgccggacc	cgggtggttc	gggtgtcgca	tgtttccat	26280
tcgccccata	tggagccgat	gctggaggag	ttcgcccaga	tctcccgagg	ccgcgaatat	26340
cacgcaccgg	aactgccgat	catctcgacc	ctgatcgggtg	agctggacgg	tggtagagtg	26400
atgggcactc	ccgagtactg	ggtgcgtcag	gtgcgtgagc	ccgtccgttt	cgccgagggg	26460
gtccaggcgc	ttgtcgggtc	gggtgtcggc	acgattgtcg	aattgggtcc	ggacggggcg	26520
ttgtcgacgt	tggtagagga	gtgtgtggcg	gaatccgggc	gggtggccgg	gatcccgctg	26580
atgcgcaagg	accgcgacga	ggcggaacc	gtgctggcag	ctttggcgca	gatccacacc	26640
cgtggtggtg	aggtggactg	gcggtagctt	ttcgccggta	ccggggcgaa	gcaagtcgac	26700
ctgcccacct	acgccttcca	gcggcagcgg	tactggctgg	catccaccgg	gcgtgcgggt	26760
gacgtgaccg	ccgccggatt	ggccgagggc	gaccatccgc	tgctcgggtg	ggtggttgcg	26820
ttggcagacg	gcgaagggtg	ggtgctgacc	ggtcgggtga	cagcgggttc	gcattcgtgg	26880
ttgtccgate	accgggtgct	ggcgaaatc	gtcgtccccg	gcaccgcgat	cgtcgagctg	26940
gtgtggcacg	tcggcgagcg	cctcggttgt	ggccgggtgg	aagaactggc	tttgggaagc	27000
cccctgatcc	tgccggatca	tggagcggtc	cagggtcagg	tgctggtggg	accgcccggg	27060
gaatccggag	cccgggtcgg	ggcgctctac	tctgtcctg	gcgaggcgat	cgaacccgag	27120
tggaagaagc	acgcgacggg	cgtgcttctc	ccaccggtgg	ccgccgagaa	ccatgagctg	27180
accgcatggc	ccccggagaa	tgcgaccgaa	atcgatgcag	acggggtcta	cgcattcctt	27240

gaagggcacg gtttcgcgta cggaccggcc tttagatgtc tgcgcgggtgc ctggcgacga 27300  
 ggcggggagg tgttcgccga agtcgcattg ccggatgaca tgcaggcggg ggtcgatcga 27360  
 ttcggcgctcc accccgcggt gctggacgcg gttctgcatg ccgccgcagc cgagacgtcg 27420  
 gtggtccaga gcgaagcgcg ggtgccgttc tctggtcggtg ggggtggaact tgcgcgcaact 27480  
 gaaagcgcgg tgggtgcgggc gcgcctctcg ttgacttcgg atgacgaact gtcgttggtc 27540  
 gcagtggacc cggctggccg attcgtggcc acggttgatt cgctggtgac ccgaccgatc 27600  
 tcccggcagc aggtgaggtc tggcgcgatc ggtgattgcc tgttcgaggt ggagtggcac 27660  
 cggaaggcgt tgttggggaac aaccgcgggc gacgacctg ccatcgtcgg tgacggtccc 27720  
 agttggccgg aatcggtgcg cgcaaccgca cggtttcgca ccctggatga gttccgtgcg 27780  
 gccgtggact cggacgttcc tgccccgggt tccgtgttg tgcagctat gtcggccgaa 27840  
 gaggtcgagg gtggatccct gccgtcgcg gcccaagagt cgacctccga tctgctggct 27900  
 ctctgacagt cgtggcttgc ggacgagcgg ttcgccgaat ccagctcgt ggtcgtcacg 27960  
 cgtgcagcgg tgtcggccga ctcgattcg gacgtcgcg acctggtggg tgcgtcgctc 28020  
 tgggggttgt tgagttcagc ccagtcggag aaccgggtc gcttcgtgct ggtggacgtg 28080  
 gacggcacac ctgagtcgtg gcaggcgttg ccggccggcc tgcgagcagg agaaccgcag 28140  
 ctggcacttc ggcgcggcgt ggcgtggtg cctcggttg cgcgactcac ggtgcgcgag 28200  
 gagggctcct ccccgcaact cgacacggac gggaccgtcc tcatcacggg tggcaccggt 28260  
 gcgttggggg gagtggttgc ccgtcacctg gtggaggagc acgggattcg gcgtttggtg 28320  
 ttggcaggcc ggcgtggtg gaatgcgcct ggagtccacg agttggtgga tgagctggcg 28380  
 cgcgcgggcg ccgtggttga ggtggtggct tgcgatgtgg ctgaccgcac cgatctggag 28440  
 cacgtgctgg ccgccattcc ggtcgactgg ccgctcgagg ggatcgtgca taccgctggg 28500  
 gtgctggccg acggagtgat cgggtccttg tccggcgccg atgtgggcac ggtgtttgcc 28560  
 ccgaaggatga cgggggcatg gcatctgcac gagttgacct gcgatctgga tctgtcgttc 28620  
 ttcgttcttt tctcttctt ctccgggatt gcgggtgccg cagggcaggc caactacgcg 28680  
 gcggcgaaaca cgttcttga tgcattggcg cgttatcgcc gggcgctgg gctgcctggg 28740  
 ttgtcgttgg cgtggggact gtgggcgcaa ccagcggta tgacgagtgg cttggacgcg 28800  
 gcgtcgggtg agcggttggc gcggacgggc atcgcagaac tttccacgga ggatggactc 28860  
 cgctgttcg atgccgcgtt cgcgaaggac cgggcttgcg tcttgccgc tcgattggac 28920  
 agggcgctgc tggtcgggaa cggacgatc cagcgcattc cggcgctgtt gagcgcgttg 28980  
 gttctgttcc gcggcggtgt ggcgaggaaa acagccaatt ctccaggccgc ggatgaggac 29040  
 gcaactgttg gtttggtgcg ggagcacgtt tccggcgtgc tgggttattc ggggtcgggtc 29100  
 gaggttgggg gcgaccgtgc tttccgtgat ctgggttttg attcgttgtc tggcgtggag 29160  
 ttgcggaacc gccttgccgg ggtgctgggg gtgcggttgc cggcgactgc ggtgttcgac 29220  
 tatccgacgc cgcgggcgtt ggcgcgttcc ctgcatcagg aactggcagg cgaggtcgcg 29280  
 tccacgtcga cgcgggtgac cagggcagcg agtgccgaag aggatcttgt tgcgattgtc 29340  
 gggatgggat gtcgttttcc ggggtggggtg tctcgcccg aggagcttg gcggtggtg 29400  
 gccggcgcg tggatgcggt ggctgggttc ccagacgatc gcggtggga tctcgcgcg 29460  
 ttgtacgatc ctgatccga tctctcggg acctcgtatg tgtgtgaggg cgggtttctg 29520  
 cgggacgcgg cggagttcga tgctgacatg ttcggcatca gcccgctga ggcgttgcg 29580  
 atggatccgc agcagcggtt gctgctggag gtcgcctgg aaaccttgga gcgggctggg 29640  
 atcgatccgt tctcgttgca cggcagccgg accggtgtgt tgcgggctt gatgtaccac 29700  
 gactatgggg ccgattcat taccagagca ccggagggtc tcgaagggca cctcgggacg 29760  
 ggcaatgcgg ggagcgtgct gtcgggtcgg gttgcgtatt cgtttggtt cgagggtcct 29820  
 gcggtgacgg tggatacggc gtgttcgtcg tctgtggtg cgttacacct ggcgggtcaa 29880  
 gcaactgcgg ccggtgagt cgaattcgcc cttgccgggt gcgtcacggt gatgtcgacg 29940  
 ccgacgacgt tctgaggatt ctcccgtaa cggggtctg ctccggatgg gcggtgcaag 30000  
 tcttccgcgg cggccgcgga tggcaccggg tggggcgagg gtgcccgtt ggtgttgctg 30060  
 gagcggttgt ccgatgccc gcgcaatgg cagcaggttc tggcgggtg gcggggtagc 30120

gcggtgaacc	aggacggcgc	gtcgaatggc	ttgactgcgc	caaatggtcc	gtcacagcaa	30180
agggatgatca	cccaggcact	cacgagtgcc	gggctgtccg	tgtccgacgt	ggatgctgtg	30240
gaggcgcatg	ggacgggcac	gcggcttggt	gatccgatcg	aggcgagggc	ggtgatcgct	30300
acgtacggcc	gggatcgtga	tcccggtcgg	ccgttggtggc	tggggtcggg	gaagtcgaat	30360
attggtcaca	cccaggcggc	ggcgggtgtc	gctgggtgtga	tcaagatggg	gatggcgatg	30420
cggcaggggg	agctgccgcg	cacgttgcac	gtggacgagc	cctccgcgca	ggtggactgg	30480
tctgcggggca	cggccaact	cctcacggag	aacacgccct	ggcccagacag	cggtcgtctt	30540
cgccggggcgg	gcgtgtcatc	gttcgggatc	agtggcacca	acgcgcacct	gaccttgaa	30600
caacctccgc	gagatcgca	gcgtcaaca	gagccggatt	cgggttctgt	ccgcgatttt	30660
ccggtggtgc	cgtggatggg	gtcgggcaaa	acaccggaag	cgtatccgc	ccaggcagat	30720
gcattgatgt	cctacttgag	caatcgcggt	gatgcttccc	cgcgagatat	cggttattcg	30780
cttgcggtga	cccgtccggc	gttgaccac	cgcgctgtcg	tgtggtgc	ggatcggtcc	30840
gcgttgctgc	cgggcttgaa	agcgtggcc	gttagtaatg	acgtgccga	ggtgatcacc	30900
ggcactcgtg	ccgtggggc	ggtcgattc	gtgttctccg	gtcaagggtg	tcagtggccc	30960
gggatgggaa	gcgggtcca	ctcggcggtt	ccggtgttcg	ccgacgcgtt	tgacgaagcc	31020
tgtcgagc	tgatgcgca	tctcgggcag	atggcccggc	tacgagatgt	gtgtccggg	31080
tcggatacgc	aacttctgga	ccagaccttg	tgggcgcagc	cgggcctgtt	cggttgcaa	31140
gtcggactct	gggagttgtt	gggttcgtgg	ggtgtccggc	ccgctgtggt	gctgggccac	31200
tcggtcggtg	agctggcggc	ggcgttcgcg	gctggagtgt	tgtcgttcg	ggatgcggct	31260
cggctggtgg	cgggccgtgc	ccggttgatg	caagccctgc	caactggcgg	tgccatgtc	31320
gctgcggctg	ctggagagga	gcagctgcgc	ccgttgctgg	ccgactgcgg	tgatcgtgtg	31380
gggatcgccg	cggccaacgc	tcccgggtcg	gtggtgctct	ccggtgatcg	ggatgtgtc	31440
gatgacattg	ccggtcggct	ggacgggcaa	gggatccggt	ccagggtggt	gcgggtttcg	31500
catgcgtttc	attcgatcg	gatggatccg	atgctggcgg	agttcaccga	aatcgcccgg	31560
agcgtggact	accggtcgtc	agggctgccg	atcgtgtcga	cgttgacggg	tgagctcgat	31620
gaggtcggca	tgccggctac	gccggagtat	tgggtgcgc	aggtgcgaga	accgtccgc	31680
ttcgccgacg	gtgttgctgc	gctcgcggct	cacggtgtga	gcaccgtcgt	cgaggtcggg	31740
ccggatgggg	tggtgtcggc	gctggtgcag	gagtgcgcgg	ccggatccga	tcaggcgga	31800
cgggtggccg	cggttccgct	catgcgcagc	aatcgcgacg	aggcgcacac	ggtgacaacg	31860
gcattggcgc	agatccatgt	gcgtgggtgct	gagggtggact	ggcggtcgtt	tttcgccggg	31920
accggggcaa	agcaggtcga	gctgcccacg	tatgccttcc	aacgacagcg	gtactggctt	31980
gactcaccat	ccgaaccggg	cgggcaatcc	gccgatcccg	cgcgcagtc	gggcttctgg	32040
gaactcgtcg	agcaggaaga	tgtcagcgcg	ctcagcgccg	ctctgcacat	taccggcgat	32100
cacgacgtgc	aggcgtccct	ggaatcgggtg	gttcgggtcc	tctcctcctg	gcacgcggg	32160
atccgcaacg	aatccctggg	gcaccagtgg	cggatccgga	tttcttgga	tgagcgggca	32220
gatttgccag	accctcgtt	gtcggggaca	tggctcgtcg	tcgtgccgga	ggggtggtcg	32280
gcgagtcggc	aagttctgcg	tttcaacgag	atgttcgagg	aacgggggtt	cccgccagtt	32340
ctgttcgagc	tcgccgggca	cgacgaggaa	gccctggcgc	aacgattccg	ctcgttgcc	32400
gttgcgtcag	ggggaataag	cggcgtgttg	tccttgctgg	cgtggatga	atcgccgtcc	32460
tcgccgaacg	ctgctttgcc	gaatggcgcg	ctgaactcgt	tggtagtctg	gcgagctctg	32520
cgggcgcggg	atgtgtcggc	gccattgtgg	ttggcgacgt	gtggtggtgt	cgcggtcggg	32580
gatgtgccgg	tgaaccggg	gcaggcgctg	gtgtggggac	tgggtcgcgt	cgtcgggtctg	32640
gagcatccgg	cctggtgggg	tggcctgggtc	gacgtgccgt	gcttgctcga	tgaggacgct	32700
cgagaacgct	tgctcgtcgt	gttggcagggt	cttggcgagg	acgagatcgc	ggtacgtccc	32760
ggtggtgtgt	tcgtgcggcg	gttggaacgc	gctggtgcgg	cgtcgggtgc	cgggtcgggtg	32820
tggcgtccctc	gggggacggg	gttggtgacg	ggtggtacgg	gcggtttggg	ggcgcatggt	32880
gcccgggtggt	tggcgggtgc	cggggctgag	catgtggtgt	tgaccagccg	tcgaggcgcg	32940
gcggctccgg	gcgctggaga	tttgccggcg	gagctggagg	cgtggggcgc	tcgggttttcg	33000

atcacggcct ggcagctggc cgatcgtgac gctttggccg aagtgttggc gaccattccg 33060  
gatgattgcc cgctgaccgc ggtgatgcat gcggcggggg tcgttgaagt cggcgacgtg 33120  
gcgtcgatgt gtttgaccga cttcgttggg gtgctgtcgg cgaaggcagg tgggtcgggc 33180  
aatctcgatg agttgctcgc cgatgtcgag ctggatgcct tcgtgctgtt ctcattccgtc 33240  
tcgggtgtgt ggggtgctgg cgggcagggc gcttatgcgg cggcgaatgc ctacttggat 33300  
gcgttggcgc agcagcgtcg ggcaaggggg ttggtgggga ctgcggttgc gtggggcccg 33360  
tgggccgggtg acggaatggc cgcaggtgaa ggcggtgcac agctgcgccg ggccggcctg 33420  
gtgccaatgg ctgcggatcg ggcgttgctg gcacttcagg gcgcattgga tcgtgacgag 33480  
acatccctgg tcgtggccga tatggcgtgg gagaggttcg ccccgtgtt cgccatgtcc 33540  
cgtcggcgtc cgctgctcga cgagctgccc gaagcacagc aggcgttggc ggatgcggag 33600  
aacaccactg atgctgcgga ctcggccgtc ccgctaccgc ggctcgccgg catggcagcc 33660  
gccgaacgcc gccgcgcgat gctggacctg gtgctggcgg aggcctcgat tgtgttggga 33720  
cacaacgggt ctgaccagt tggccccgac cgggcgttcc aggagctcgg atttgattcg 33780  
ctgatggccg tcgaactgcg caacaggttg ggcgaggcaa caggattgag tctgccggcc 33840  
acgttgatct tcgattatcc gagcccatcc gcgctggctg agcagctggt cggcgagctg 33900  
gtgggagcgc agcccgcgac caccgtcgtg gccggggccg atccagtgga tgatccggtt 33960  
gtcgtggctg cgatgggatg ccggtatccg ggcgacgtct gctcgcccga ggagctgtgg 34020  
cagctggttt ctgcgggacg tgatgcggta tcgacgttcc ccgtcgatcg gggttgggac 34080  
tgcaacacgt tgctcgaccc ggatccggat cgggcaggca gtacctatgt gcgagaaggt 34140  
gccttctcga ccggtgctga tcggttcgac gccgggttct tcggcatcag cctcgcgag 34200  
gcgcgcgcaa tggatccgca gcagaggttg ttgctcgaag tggcgtggga ggttttcgaa 34260  
cgagcaggaa tcgctccgct gtcgttgccg ggtagcagga ccggtgtgtt cgcggggacc 34320  
aatgggcagg accacggtgc gaaagtggct gccgcgccgg aggcggcggg tcacctcctg 34380  
accggaaacg ccgcgagtg cctggccggc cggttttctt acacgttcgg ccttgagggg 34440  
cctgcgggtg cggtggatac cgcgtgttcg tcgtcgttgg tggcgttgca tttggcgtgc 34500  
cagtcgctgc gttcgggtga gtgtgatatg gcgttggcag gtggtgtgac ggtgatgtcg 34560  
acacccctgg ctttctcga gttctctcgt cagcgcggtt tggcgccaga tggtcggtgc 34620  
aagtcgtttg cggccgctgc ggatggcacc ggggtgggtg aggtgtccgg cctggtgttg 34680  
ctggagcggg tgctcgatgc tcgtcggaat ggtcaccggg tgttggccgt ggttcgcggg 34740  
tctgcggtga atcaggatgg tgcctcgaat ggcctgactg cgccgaatgg tccgtcgag 34800  
cagcgggtga ttcggcaggc cctcgcgaat gcggggctgt cggcgtccga tgtgatgtc 34860  
gtggaggcgc acgggaccgg taccgggctc ggggatccga tcgaggcgca ggcgtgatc 34920  
gcgacatatg ggcaggagcg ggatcctgag cgggccctgt ggctggggtc gatcaagtcc 34980  
aacatcggcc acacgcaggc ggcggccggt gtggcggggg tcatcaagat ggtgcaggcc 35040  
atgcggcacg gggagttgcc tgcgacgttg cacgtggaca agcccactcc acaggtggac 35100  
tggctcgcg gggcgttcg gtcctcacc ggaacacgc cctggcccga gagcggccgt 35160  
cctcgtcgag cgggggtgtc gtcgttcggg atcagcggca ccaacgcaca cctcatcctc 35220  
gaacaaccac cgtcgggaacc agcggagatc gaccaatcgg atcggcgggt cactgcgcat 35280  
ccagcgggtg tcccgtggat gttgtcggct aggagtctcg cagcgtcga gccccaagcg 35340  
gctgcgtgc aggcgggct ggaccgggt cctggcgctt ctccgtgga tttggggtat 35400  
tcaactcgca ccaactgttc tgtgttgga gaacgcgcg tcgtgtgggg tgccgatcgg 35460  
gaggcactgc tgtccaggct ggcagcgtc gccgatggc ggacggcgcc gggggtgata 35520  
acgggctctg cgaattccgg tggccgcac ggattcgtt tttccggtca gggcagtcag 35580  
tggctgggga tgggaaaggc gttgtgcgc gctttccgg cgttcgcgga cgccttcgag 35640  
gaagcctgcg acgcgctaag cgcacacctg ggcgcggacg ttcggggtgt gctgttcggt 35700  
gctgatgagc agatgctcga ccggacgtg tgggcgcagt cggggatctt cgcggttcaa 35760  
gtcggcctcc tgggattgct gaggtcgtgg ggcgtgcggc cggccgcggt gctggggcac 35820  
tcggtcggcg agttggctgc ggcgcacgcg gctggtgtgt tgccttgcc ggacgtgca 35880

cggttggttg	cggttcgggc	ccacctgatg	caggcattgc	ccaccggcgg	cgcaatgctc	35940
gcggtcgcca	ccagcgaggc	ggcggtcgga	ccgctgcttt	ccggggtgtg	cgatcgggtc	36000
agcatcgctg	cgatcaacgg	ccccgagtcg	gtagtgtctt	ccggcgaccg	cgatgtgctc	36060
gtggagctcg	caggcgaatt	cgatgcccga	gggcttagga	ccaaatggtt	gcgggtctcc	36120
catgctttcc	actcgcaccg	gatggaaccg	attctggacg	agtacgcgga	aaccgccagg	36180
tgcgtcgagt	tcggtgaacc	ggtggtgccg	atcgtctccg	ccgcgaccgg	tgcgtggac	36240
accaccggac	tgatgtgcgc	ggccgactac	tggacgcgcc	aagtgcgtga	tcctgtccgc	36300
ttcggagacg	gtgtccgggc	gctcgtcggc	caaggcgtgg	acacgatcgt	cgagttcggc	36360
ccggacgggg	cgttgtcggc	cctggtcgag	cagtgtcttg	ccgggtccga	ccaggctggg	36420
aggggtggcg	cgatcccgtt	gatgcgcagg	gaccgcgatg	aggtcgagac	cgcggtggcg	36480
gccctggcgc	acgtgcacgt	ccgcggtggt	gcggtggact	ggtcggcttg	cttcgccggc	36540
accggcgccc	gcaccgtcga	gttgcccacc	tacgccttcc	aacgccagcg	gtactggctg	36600
gccgggcaag	cggacggggc	cggcggcgat	gtggttgccg	acccggtcga	cgcgcttcc	36660
tgggagttgg	tcgagcgcgc	cgatccggaa	ccgttggtgg	atgaactctg	catcgaccgg	36720
gaccagccct	tccgggaggt	gctgcccgtt	ctggcttcc	ggcgcgagaa	acaacgccag	36780
gaggccctcg	cggattcctg	gcgctaccag	gtgcgtgga	ggtccgtcga	ggtgccgtcc	36840
gcagccgccc	tccggggcgt	gtggctggtg	gtgcttccag	ctgacgtgcc	ccgagatcaa	36900
ccggcggtcg	tcacgcacgc	gctgatcgcg	cgcggcgcgc	aggtcgcggt	cctggaattg	36960
accgagcagg	acctccaacg	cagtgcgctt	gtggacaagg	tgcgcgccgt	cattgcggac	37020
cgcaccgagg	tgacgggtgt	gttgtctctg	ttggcgatgg	acggcatgcc	ctgcgcggcg	37080
catccgcacc	tgtcccgtgg	tgtcgccgct	accgtgatcc	tgacgcaggt	gttgggcat	37140
gcgggtgttt	ccgccccgct	gtggttgccc	acgaccggtg	gcgtcgaggc	cgggaccgag	37200
gacggtccgg	ccgatccgga	ccacggcttg	atctgggggc	tcggcagggg	cgtcggcctt	37260
gaacatccgc	agtgtggggg	tggcctgac	gaccttccgg	agacactgga	cgagacgtcc	37320
cggaaacgggt	tgggtggccg	actcgccggg	acggcgccgc	aagatcagct	cgccgtgcgt	37380
tcacccgggt	tgttcgttcg	cagagtgggt	cgcgacgcgc	ggaacccccg	gtcagagaca	37440
tggcgtagcc	ggggaacggt	cctcatcacg	ggcggaacag	gcgcgctcgg	tgccgaggtc	37500
gcacgatggc	tggcccggcg	gggagctgag	cacctggtgt	tgatcagtcg	ccgcggcccg	37560
gaagctcccg	gcgcagcgga	cctagggggc	gagctgactg	aactcggcgt	gaaagtcaca	37620
gtcttgccct	gcgatgtgac	ggaccgcgac	gagctggcgg	cggtgctggc	ggccgttccc	37680
acggagtatc	cgctgtcggc	ggtcgtgcac	accgccggcg	tcgggacgcc	tgcgaacctg	37740
gccgagacga	ccttggcgca	gttcgccgac	gtgttgctcg	ccaaggctgt	cggcgcggcg	37800
aacctggacc	ggctgcttgg	cgggcaaccg	ttggacgcct	tcgtgctggt	ctcctcgatc	37860
ccgggagttt	ggggagccgg	cggccaagga	gcctattcgg	ccgccaatgc	gtatctcgat	37920
gcccttgccg	agcgccgacg	ggcttgccgg	cggccggcga	cgtgcacgcg	ctggggtccg	37980
tgggcccgggt	cgggcatggc	cgttcaggaa	ggtaacgagg	cgcactctcc	ccgaaggggc	38040
ctggtaccga	tggaaaccga	gtcggccctc	ttcgcgctgc	aacaggccct	gtcccaacga	38100
gaaaccgcca	tcaccgtcgc	agatgtggac	tgggagcgat	tcgccgcctc	tttcaccgcg	38160
gcccgcgccg	gaccactggt	ggaagagatc	gtggatctac	ggcccgcacac	cgagaccgag	38220
gagaagcacg	gtgccggcga	gctggggcag	cagctggccg	cactgccgcc	cgtgagcgc	38280
ggacacctgc	tgctggaggt	ggtgctggcg	gaaaccgcca	gcaccctggg	gcacgattcg	38340
gcggaggctg	tgcaaccgga	tcggaccttc	gccgaactgg	gcttcgattc	gctgaccgcg	38400
gtagagctgc	gcaacaggtt	gaacgcgggt	accgggcttc	gcctgccgcc	gacgctggtt	38460
ttcgaccacc	cgacgccgct	ggcgttgctc	gaacagttgg	ttccggccct	ggtcgcggag	38520
ccggacaacg	gcatcgaatc	gctgctcgcc	gagctcgaca	ggctggatac	cacgttggcg	38580
caagggcctt	cgatcccact	ggaagaccag	gccaaagggt	cggagcgctt	gcacgcactc	38640
ctcgccaagt	gggacggggc	gcgtgacggc	acggccagag	cgacgtcacc	ccaatcgctg	38700
acggcggccca	cggacgacga	aatcttcgac	ctcatcgacc	ggaagtcccg	gcgctgaccg	38760

ccccttctc gctcagctc cctgattac tggaaacggtg tatttcgatg gccaatgaag 38820  
 aaaagctccg cgagtacctc aagcgtgtcg tcgtcgaact ggaagaggcg cacgaacgcc 38880  
 tgcacgagtt ggagcgccag gagcacgacc ccatacgcat cgtgtcgtat ggatgtcgtt 38940  
 atcccgggtg cgtctccact ccggaggagc tgtggcgact ggtcgtcgac ggaggagacg 39000  
 cgatcgcgaa cttccccgaa gaccgtggct ggaatctgga cgagctgttc gatcctgatc 39060  
 cgggcccagc cgggacctcc tacgtccgag aggggtggtt cctgcgcggg gtcgcggact 39120  
 tcgatgccgg gctcttcggg atcagtccgc gcgaggcaca ggcgatggac ccgcaacagc 39180  
 ggttgctgct ggagatctcg tgggaggtgt tcgagcgcg cggcattgac ccgttttctt 39240  
 tgcggggtac caagaccggt gtgttcgcgg gcctgatcta ccacgactac gcgtcgcgg 39300  
 ttcgcaagac cccgcggag ttcgaggggt acttcgccac cggcaacgcg ggcagcgtcg 39360  
 catccggccg ggtggcttac accttcgggt tagagggcc ggcggtcacc gtggacaccg 39420  
 cctgctcgtc gtccttgggt gcgtgcacc tggcctgccat gtccctgcgg ctgggcgaat 39480  
 gcgacctggc cctggccggg ggcatttcgg tgatggccac gccgggagcc ttcgtcgagt 39540  
 tcagccggca acgcgcactc gcctcggatg gccggtgcaa gcccttcgcg gatgccgccg 39600  
 acggcaccgg ctggggcgag ggcgcggaa tgctgctgct ggaacggctg tcggacgcac 39660  
 gacgaaacgg ccaccgggtg ctggcgccgg tggtcggttc cgcgatcaac caggacggga 39720  
 cgtccaacgg cctgaccgcg cccagcggtc ccgcacagca gcgagtgatc cgccaagccc 39780  
 tggcgaacgc cgggttgctg cccgcggagg tcgatgtggt cgaggcgac ggcacgggca 39840  
 cggccttggg cgacccgatc gaggcgcagg cctgatcgc cactacggg gcgaaccggt 39900  
 cggcggatca tccgtgctg ctgggttccc tcaagtcgaa catcgccac acccaggctg 39960  
 ccgcccgtgt ggcgggggtg atcaagtcgg tccctggccat caggcaccgg gagatgccc 40020  
 gcagcctgca catcgaccag ccatacgagc acgtggactg gtcggcgggc gcggtgcggc 40080  
 tgctcacgga cagcgttgac tggccggatc tcggcaggcc gcgcgagca ggggtgtcct 40140  
 cgttcggcat gagcgttacc aacgcacacc tgatcgtcga ggaagtatcc gacgagcccg 40200  
 tctcgggcag taccgagccg accggggcat ttccttggcc gctgtccggc aagacggaga 40260  
 cggcattgcg cgagcaggct gccgagttgc tctccgtagt gaccgagcac ccggagcccg 40320  
 gactggggga cgtcgggtac tcgctggcca ccggtcgcgc tcgatggag caccgggctg 40380  
 tcgtggttgc cgacgatcgg gactctttcg tcgccggact gacggcggtg gctgcgggag 40440  
 ttcgggcagc caacgtggtg cagggcgcgg ccgactgcaa gggaaaggct gcgttcgtgt 40500  
 tccccggcca gggctcgcgt tggcagggga tggcgagggga actgtccgaa tccctgcgg 40560  
 tgttcggcg gaagctggcg gaatgcgcgg cggctacggc cccttacgtg gactggtcgc 40620  
 tgctcggcgt ccttcgcggg gatcccgatg caccgcgct ggatcgcgac gacgtgattc 40680  
 agctcgcgt gtccgccatg atggtgtcgc tggccgaact gtggcggttcg tcgggagtg 40740  
 agccgcgcg ggtggtcggg cattcccagg gcgagatcgc cgccgcccgt gtggcaggcg 40800  
 ctttgtcctt gactgatgag gtgcgcacat tcgctgcccg ctgcgatgag gtgtcggcgc 40860  
 tgaccgggaa gggaggcatg ctgcgattg ccttgccgga aagcgcggtg gtgaagcgaa 40920  
 tcgaggcct gccggagctg accgttgcgg cggcacaagg acccggtcc actgtcgtt 40980  
 ccggcgaacc gtcggctctg gagcgtctgc agaccgaact gaccgcggaa aacgtgcaga 41040  
 cccggcgggt gggaaattgat tacgcctcgc attgcgcga gatcgcgag gtccagggcc 41100  
 ggcttctgga ccggctgggc gaagtcgggt ccgaacctgc tgagatcgt ttctactcga 41160  
 cggtcaccgg cgagcggagc gacaccggcc gactcgacgc cgactactgg taccagaacc 41220  
 ttcggcagcc cgtccgcttc cagcagaccg tcgcccgat ggcagatcag ggctatcgg 41280  
 tcttcgtcga ggtgagccc caccgctgc tcaccgcgg aatccaggaa acgctggaag 41340  
 ccgcggacgc gggcgggggt gtggtcgggt cgtgcggcg tggcgagggc ggctcccggc 41400  
 gctggctgac ttcgctggcc gactgccagg tcgcgggact gccggtgaat tgggaacagg 41460  
 tattcctcaa caccggagcc cgacgcgtgc cgtgcgcac ctaccgcttc cagcggcagc 41520  
 ggtactggtt ggagtcgcgc gactacgag cggcgatct cggttcgggt ggcttgcct 41580  
 ccgcgagca tcccctgctc ggggtgcgg tgacgctggc cgatgcgggc gggttcctgc 41640

tgaccggcaa	gctgtcggtc	aagaccacgc	cctgggtggc	cgaccacgtg	gtcggcgggg	41700
cgatcctgct	gcccggcacc	gcgttcgtgg	aaatgctgat	acgcgccgcg	gaccaggtcg	41760
ggtgcgatct	gatcgaggag	ttgtccctga	cgactccgct	ggttttgccc	gcgaccggtg	41820
cgggtgcagg	gcagatcgcg	gttggcggtc	cggacgaggc	cgggcgccgc	tcggtccgcg	41880
tgcattcctg	tcgagacgac	gccgtgccgc	aggactcgtg	gacctgccac	gcgaccggca	41940
cgttgacctc	cagcgatcac	caggacgccg	gccagggccc	cgatgggatt	tggccgccc	42000
acgatgctgt	cgcggttccg	ctggacagct	tctacgcccg	cgcagctgag	cggggcttcg	42060
atttcggccc	ggcgttccag	gggttgacag	cggcttgga	gcgcggagac	gagatcttcg	42120
ccgaggtcgg	cctgcccacc	gcacaccgcg	aagacgcccg	caggttcgga	atccaccctg	42180
ctctgctgga	tgcggcactg	caggcgctgg	gcgcagccga	agaggatccg	gacgagggat	42240
ggctcccgtt	cgcgtggcaa	ggtgtgtccc	tcaaagcgac	gggcgcactt	tcccttcggg	42300
tgcacctcgt	tccggcgggc	gcgaatgcgg	tgtcgggtgt	cacgaccgac	acgactggcc	42360
aagccgtgct	ctccatcgat	tcgctgggtg	tgcgccagat	ttcggacaag	cagttggcag	42420
cggcccgtgc	gatggaacac	gagtcctcgt	tccgggtcga	ctggaagcga	atctcgccc	42480
gcgctgccaa	gccggtctcc	tgggcagtg	tcggcaatga	cgaactcgcc	cgagcctgcg	42540
gctcggcact	tggcacggaa	ctccaccccg	acctgaccgg	gttggtgac	ccgccccggg	42600
acgtcgtggt	ggtgccatgc	ggtgcgtctc	gccaggactt	ggacgttgct	tccgaggcac	42660
gtgcgcgcg	acaacgcatt	cttgacctga	tccaggattg	gttggcggcg	gcgcgattcg	42720
ccggatctcg	cctggtgggt	gtgacgtgtg	gtgcggcgct	gacaggtccc	gccgagggtg	42780
tttcgcacct	ggtgcattgc	gcgtcgtggg	gtttgttgcg	ttcggcgag	tcggagaacc	42840
cggaccgatt	cgtgttggtc	gatgtggacg	gaaccgccga	atcatggcgt	gcgctcgcg	42900
cggcgcgtgc	ttccggagaa	ccgcagctgg	cgttgcgcg	cgggtgaagtc	cgggtgcctc	42960
gcctggcgcg	atgtgttgcc	gccgaggaca	gccggatccc	agtccccggt	gcggatggga	43020
cgggtgtgat	ttccggcggt	acgggcctgc	tgggcgggtt	ggttgcccgg	catttggtgg	43080
cggagcgcg	tgtccgccgc	ctggtgctcg	cggggcgacg	cggctggagc	gcccccgggg	43140
tcaccgacct	ggtggatgag	ttggtggg	tgggagctgc	ggtcgagggt	gcgagctgcg	43200
atgtcgggga	tcgggcccag	ttggaccggc	tgtgacgac	gatctcggca	gagttcccgc	43260
tgcgcggagt	ggtgcattgc	gccggggcac	ttgccgacgg	ggtcgtcgag	tcgctgacac	43320
cagagcacgt	ggcaaagggt	ttcggcccga	aggccgccgg	tgcgtggcac	ctgcacgagt	43380
tgactcttga	tctggatctc	tcgttcttcg	tgtctctctc	ctcgttctcc	ggcgtggcgg	43440
gggctgcggg	tcagggaaac	tacgcggcgg	cgaacgcgtt	cctggacggc	ctggctcagc	43500
accggcggac	ggcggggctg	cctgcgggtg	cgtcggcttg	gggcttggtg	gagcagccca	43560
gcgggatgac	cggagcgctc	gatgcggcgg	gccgtagccg	cattgcgcgc	accaatccgc	43620
cgatgtccgc	gccggacggg	ttgcggctgt	tcgagatggc	gtttcgcgtt	ccgggcgaat	43680
cgtctctggt	tccggtccac	gtcgacctga	acgccctgcg	cgtgatgcg	gccgacggcg	43740
gtgtgcctgc	gttgttgccg	gacctgggtg	cagcgcccg	gcggcggagc	gcggtcaacg	43800
agtcggcgga	cgtcaacggg	ctggttggtc	ggctgcggag	gctgccggac	ctggatcagg	43860
aaaccagct	gttggttttg	gtgcgcgagc	atgtttcggc	ggtgctgggg	cattcgggtg	43920
cggtcgagg	cggggccgat	cgtgctttcc	gggatttggt	ttttgattcg	ttgtccggtg	43980
tggagtttcg	gaaccggctt	ggcgggggtg	tgggcgttcg	gttgccggct	actgcgggtg	44040
tcgactatcc	gacaccgcgg	gcgttggttc	ggttcttgct	cgacaaactg	attggtggcg	44100
tggaggctcc	gactcccgc	ccggcggtcg	tggcggcgg	gactgctgac	gaccccggtg	44160
tgatcgtggg	gatgggctgt	cgttatccgg	gtgggggtgc	ctcgccggag	gagctttggc	44220
gtttgggtgg	cgggggcttg	gatgcgggtg	cggagttccc	ggacgatcgt	ggctgggac	44280
aggcgggggt	gttcgatccg	gatcccgatc	gtcttgggac	ctcgtatgtg	tgtgagggtg	44340
gcttctcgcg	agatgcggca	gagttcgatg	ccggtttctt	cgggatttcc	ccgcgtgagg	44400
cgttggcgat	ggatccgcag	cagcgggtgc	tgttggaagt	cgttgggaa	accgtggagc	44460
gggcggggat	tgatccgctt	tcgttgccgg	ggagccggac	cggcgtgttc	gcggggctga	44520



tgcaccacga ctacggcgcg cggttcatca cgagggcgcc ggagggtttc gagggttatc 44580  
 taggtaatgg cagcgcgga ggcgtgtttt cgggtcgggt tgcgtattcg tttggtttcg 44640  
 agggtcctgc ggtgacggtg gatacggcgt gttcgtcgtc gttggtggcg ctgcacctgg 44700  
 cgggtcaagc actgcggtct ggtgagtgtg atctggctct tgcgggtggt gtgacggtga 44760  
 tggccacgcc ggggatgttc gtggagtttt cgcgtcaacg gggcttggcg gcggatgggc 44820  
 ggtgcaagtc gtttgcggcg gctgcggatg gcaccgggtt gggagaaggc gcgggcttgg 44880  
 tgttgttga gcggtgtcg gatgcccggc gcaacgggca cgcggttctg gcggtcgtgc 44940  
 ggggtagcgc ggtgaatcag gatggtgcgt cgaatggttt gacggcgccg aatggggcct 45000  
 cgcagcagcg ggtgatcacg caggcgttgg cgagtgtgtg tttgtcgggt tctgatgtgg 45060  
 acgccgtgga ggcgcattgg actggaacca ggcttgggtga tccgattgag gcgcaggctc 45120  
 tgattgccac ttacgggcag gggcgggata gcgatcggcc gttgtggtt gggtcgggtga 45180  
 agtcgaatat tggtcatacg caggcggcgg cgggtgtcgc tgggtgtgatc aagatggtga 45240  
 tggcgatgcg gcacgggcag ctgcccgca cgttgcattg ggatgaacct acgtcgggaag 45300  
 tggattggtc ggcgggggat gtccagctcc tcacggagaa cacccttgg cccggcaaca 45360  
 gccatcctcg gcgggtgggc gtgtcgtcgt tcgggatcag cggcaccaac gcacacgtca 45420  
 tcctcgaaca agcctcgaaa acaccagacg agactgcgga caagagcggc cccgattcgg 45480  
 aatcgaccgt ggaccttcca gcggtcccgt tgatcgtgtc ggggagaaca ccggcagcgc-45540  
 tcacgcgtca ggcgagcgca ttgttgcct atttgggtga gcgtggcgat atttccacgc 45600  
 tggatgcggc gttttcgtt gcttctccc gggccgcgtt ggaggagcgg gcggtggtgc 45660  
 tgggagcggc ccgcgaaacg ttgttgcct ggttggaaagc gctggcttcc ggtcgcgagg 45720  
 cttctggggg ggtgtcggga tccccggtct ctggcggggt tgggttcgtg ttcgccggtc 45780  
 agggcggaaca gtggttggg atgggdcggg ggctctactc ggtttttccg gtgttcgctg 45840  
 acgcgtttga cgaagcatgt gccggactgg acgcgcattc ggggcaggac gtgggggtcc 45900  
 gggatgtggt gtttggttcc gacgggtcct tgttggatcg gacgctgtgg gccagtcgg 45960  
 gtttgttcgc gttgcaggtt gtttgcgtga gcctgctggg ttcgtgggggt gtccggccgg 46020  
 gtgtggtgct gggccattcg gtcggcgagt tcgcggcggc ggttgcggcg ggagtgtgt 46080  
 cgttgcggga tgcggctcgg atggtggcg gtcgtgcccg gttgatgcag gcgttgcctt 46140  
 ctggcggtgc catgttggcg gtggtgctg gtgaggagca gctgcggccg ttgttggccg 46200  
 atcgggttga tgggtcgggt atcgcgcgg tcaacgctcc tgagtccgtg gtgctctccg 46260  
 gcgatcggga ggtgcttgac gacatcgccg gcgcgctgga tgggcaaggg attcgggtggc 46320  
 ggcggttgcg ggtttcgcct gcgtttcatt cgtatcggat ggaccgatg ttgcaggagt 46380  
 tcgccgaaat cgcacgcagc gtggactacc ggcgtggcga cctaccggtc gtgtcgacgt 46440  
 tgacgggtga gtcgacacc gcagggtgtga tggctacgcc ggagtattgg gtgcgtcagg 46500  
 ttcgagagcc cgtccgcttc gccgacggcg tccgggtgct cgcgcagcaa ggggtcgcca 46560  
 cgatcttcga actcggccct gatgcgacgc tgtcggccct gattcccgat tgtcattcgt 46620  
 gggctgatca ggccatgccg attccgatgc tgcgtaaaga ccgtacggaa accgaaactg 46680  
 tggtcgccgc ggtggcgcg gcgcacacgc gtggtgttcc ggtcgaatgg tcggcgattt 46740  
 tcgccggcac cggggcacgg cgggtcgagt tgcgcagcta tgccttcag cggcagcgg 46800  
 actggctgga aacatcggt tacggcgatg tgacgggtat cggcctggct gcggcgaggc 46860  
 atccgttgct gggggcgtg gttgcgtgg ccgatggtga tgggatggtg ctgaccggcc 46920  
 ggttgcgtg ggggacgcct ccgtggctgg cccagcatcg cgtgctgggc gaggtcgtcg 46980  
 tccccggcac cgccatcctg gagatggccc tgcacgcagg ggcgcgtctc ggctgtgacc 47040  
 ggggtggaaga gctcaccctg gaaacaccgc tgggtgtccc cgaacgcgcg gcgggtgccc 47100  
 gtagtcgtgg cctgcggga gggaccacag tttcaattga aactgcggaa gaacgtgtgc 47160  
 ggacgaacga cgccatcgaa atccagctgc tgggtgaacgc acccgacgaa ggcggtcggc 47220  
 gaagggtgtc gctgtattcc cgcgcggccg gtgggtcgag aggtgggggt tggacgcgcc 47280  
 acgccaccgg cgaactcgtc gtcggcacca ccggtggtag ggcggttccct gattggtcgg 47340  
 ctgagggtgc cgagtcgatt gctctcgatg agttctacgt cgctctggcc ggaaacgggt 47400



tcgagtacgg gccgttggtc caggggcttc aggcggcatg gcgtcgtggt gacgaggttc 47460  
 tcgccgaat cgccccgccg gccgaggccg atgcgatggc gtcgggatac ctgctcgacc 47520  
 cagcgttgct ggatgcccg ctcgaggcgt ccgcgctcgg cgaccgcccg gagcaaggcg 47580  
 gcgcgtggct gccgttctca ttcaccggcg tcgaactttc cgctccggca gggacgatca 47640  
 gcaggggtcg gctggagacc aggcgaccgg acgcgatata ggtggccgtg atggatgaga 47700  
 gtgggcggtt gctcgccctc atcgattctc tcaggctacg aagcgtgtcg tcgggacagc 47760  
 tggcgaatcg ggacgctgtc cgcgacgcgc tgttcgaggt gacctgggag ccggtggcga 47820  
 cgcagtcgac ggaaccgggt cgctgggccc tgcttggtga tactgcctgc ggtaaagacg 47880  
 atctcatcaa actcgcaacg gattccgccg accgctgcgc ggatctggcg gcgctagccg 47940  
 agaaacttga ttccagcgcg ctggttcctg atgtcgtggt ctactgcgcc ggagaacagg 48000  
 cggatcccgg caccggcgca gccgcacttg cggagaccca gcagacgttg gctctgctcc 48060  
 aagcgtgggt ggctgagccg cggttggccg aggcacgtct ggtggtggtg acgtgtgcag 48120  
 cggtgacgac ggctccgagt gacggtgcat cagagctggc acatgcgccg ttgtgggggt 48180  
 tgttgctgctc cgcgcaggtg gagaaccggg ggcagtttgt gctggcggac gtcgacggaa 48240  
 ccgccgaatc gtggcgtgct ttgccgagt cgttgggctc gatggaaccg cagttggccc 48300  
 tgccgaagg cgcggtgcga gcgccccgt tggcttcggt cgccgggcag atcgacgtgc 48360  
 ccgcggttgt ggcgatccc gaccgaaccg tgctgatttc gggcggcacg ggcctgttg 48420  
 ggggcgcggt tgcccgccac ctggtgaccg aacgcggtgt ccgccgattg gtgttgacgg 48480  
 gccgtcgtgg ctgggatgct cctggaatca ccgagttggt gggtagctg aacggcctcg 48540  
 gtgccgtggt cgacgtggtg gcgtgcgacg tcgcgatcg tgctgatctg gagtgcgttc 48600  
 tggcggcggt ccgcgggaa tttccgttgt gcggcggtg gcacgcccg ggggcgtggt 48660  
 ccgacggggt gatcgagtcg ttgtcaccgg acgacgtggg agcgggtgtt gggccgaagg 48720  
 cggcgggggc gtggaatctg cacgagctga ctcgtgatac ggacctgtcg ttcttcgct 48780  
 tgttctctc gctttccggt gttgccggcg ctcttggtca gggcaattat gcggcgcgca 48840  
 acgcgttctt ggacgcattg gcgcattacc ggcggtcaca gggactgcct gcggtgtcgc 48900  
 tggcctgggg cctgtgggag cagccgagcg ggatgacgga gacgctcagc gaggtcgacc 48960  
 ggagcaggat cgcgcgcgcc aaccgcgct tgtccacca ggagggttg cggctgttcg 49020  
 atgccgggct ggcgctggac cgggcagcgg tagttccggc gaagtggac aggactttcc 49080  
 tggccgagca ggcgcggtcg ggctcgctgc ccgcattgtt gacggcactg gtaccccca 49140  
 tccgtcgtaa taggcgggct agcggaaacc agctcgcgga cgagggcacc ctgctcgggg 49200  
 tggtgcggga gcatgcccg gccgtgctgg ggtattcgag cgcggctgac gtcggggtcg 49260  
 agcgcgcttt ccgggatctg ggttttgatt cgttgtctgg tgtggagttg ccgaaccgcc 49320  
 ttgccggggt gctgggggtg cggttgccgg cgactgcggt gttcgactat ccgacgccga 49380  
 gggcgctggc ccggttcctg caccaggaac tggcagacga gatcgctacg acgccagcgc 49440  
 cggtgacgac gaccagggca ccggtcgccg aagacgatct cgtcgcgata gtcgggatgg 49500  
 gatgccgttt tcccggtcag gtgtcctcgc cggaggagct ctggcggttg gtggccgggg 49560  
 gcgtggatgc ggtcgcggac ttcccagccg atcgcggtg ggatctggca ggcttgctcg 49620  
 atccggaccc ggaacgggct gggaaagacct acgtgcggga aggggccttc ctaccgacg 49680  
 ccgatcggtt cgatgcgggt ttcttcggga tttccccgcg tgaggcggtg gcgatggatc 49740  
 cgcagcaacg gctgttgctg gagctgtcct gggaggccat tgaacgggca gggatcgatc 49800  
 cgggttcgct gagggggagt cggaccgggtg tgttcgcggg gctgatgtac cagactatg 49860  
 gcgcccgggt cgccagccga gcccgggaag gtttcgagg gtatctcggc aatggcagtg 49920  
 ctgggagtg cgcgtcgggc cggattgcgt actcgtttgg ttccgagggt cctgcggtga 49980  
 cggtggtatc tgcgtgttcg 50000

<210> 6

<211> 29736

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 6

tcgtcggttg tggcgttgca tttggcgggt cagtcgttgc gttccggcga atgcgatctc 60  
gcccttgccg gtggtgtgac ggtgatgtcg acgcccggga cgtttgtgga attctcccgt 120  
cagcggggcc tggcaccgga cgggcgggtgc aagtcgttcg cggagagcgc ggacgggtacc 180  
ggttgggggtg aggtgtgctg tttggtgttg ttggagcggg tgtcggatgc tcggcggaat 240  
gggcatcggg tgttggcggg ggttcgtggg tcggcgggtga atcaggatgg tgcgtcgaat 300  
ggcttgaccg cgcgcaatgg tccctcgcag cagcgggtca tccagcaggc gttggcgagt 360  
gcgggtctgt cgggtgtccga tgtggatgcc gtggaggcgc atgggaccgg gaccaggttg 420  
ggtgatccga ttgaggcgca ggctctgatt gctacgtatg ggcgcgatcg tgatcccggg 480  
cggccgttgt ggttgggggtc ggtgaagtcc aacatcgggc atacgcaggc ggcgcggggt 540  
gttgccgggtg tgatcaagat ggtgatggcg atgcggcacg ggcaacttcc gcgcacgctg 600  
cacgtggatg caccctctc gcaggtggat tggtcggcgg ggagggtcca gtccttgacg 660  
gagaacacgc cctggcccga cagtggctgc ccctgtcggg tgggggtgtc gtcgttcggg 720  
atcagcggca ccaacgcgca cgtcatcctg gaacagtcca cggggcagat ggatcaggca 780  
gcggagccgg attcagatcc tgttctggat gttccgggtg tgccgtgggt ggtgtcgggc 840  
aaaacacccg aagcgctatc cgcccaggcg gcaacgttgg cgacctattt ggacaaaaat 900  
gttgatgtct cccctctgga cgttgggatt tcgcttgccg tgacctgttc ggcgctggat 960  
gagcggggcg tgggtgtggg gtcggatcgt gacacgttgt tgtctggcct gaatgcgctg 1020  
gctgccggtc atgaggctgc tggcgtgggt acgggacctg tcgggattgg tggccggacc 1080  
gggtttgtgt tcgccggtea aggcggtcag tggttgggga tgggcccgcg gttgtactcg 1140  
gagtttccgg cgttcgccgg tgctttcgac gaagcatcgc ccgagctcga tgcgaacctg 1200  
gggagggaag tcgggggttcg ggatgtgggt ttcggctccg acgagtcctt gctggatcgg 1260  
actttgtggg cgcagtcggg tttgttcgcg ttgcaggctg gtctctggga attgttgggt 1320  
acgtgggggtg ttcggcccag cgtagtgtcg gggcattcgg tcggggagct agccgcggcg 1380  
ttcgccgcag gtgtgctgtc gatggcggag gcggctcggc tgggtggcggg tcgtgcgcgg 1440  
ttgatgcagg cgttgccttc tggcgggtgcc atgctggcgg tgtccgcgac cgaggcccga 1500  
gtcggcccgc tgctcgatgg ggtgcgggat cgtgttgggt tcgcagcggg taacgctccg 1560  
gggtcgggtg tgctttccgg tgaccgggat gtgctcgatg gcattgccgg tcggctggac 1620  
gggcaaggta tccggtcgag gtggttgccg gtttcgcacg cgtttcattc gcatcggatg 1680  
gatccgatgc tggcggaggt cgcgcgagtc gcacggagcg tggactaccg gtctccacgg 1740  
tgccgattg tctcgacgct gaccggaaac ctcgatgacg tgggcgtgat ggctacgccg 1800  
gagtattggg tgcgccaggt gcgagagccc gtccgcttcg ccgacggtgt ccaggcgctt 1860  
gtggaccaag gcgtcgacac gattgtggaa ctcgggtccg acggggcggt gtcgagcttg 1920  
gttcaagagt gtgtggcgga gtccgggcgg gcgacgggga ttccgttggt gcggagagac 1980  
cgtgatgagg tccgaacggt gctggacgct ttggcgcaga cccacactcg tgggtggcgcg 2040  
gtggactggg ggtcattttt cgctggtacg agggcaacgc aagtcgacct tcccacgtat 2100  
gccttccaac gacagcggtg ctggctggag ccatcggatt ccggtgatgt gaccgggtgtt 2160  
ggcctgaccg gggcggagca tccgctgttg ggtgccgtgg tgccggtcgc gggcgcgcat 2220  
gaggtgctgc tgaccggcag gctgtcgggt gggacgcata cgtggctggc ggaacaccgc 2280  
gtgctggggc aagtcgtcgt ccccgccacc gcgttgctgg agatggcgtg gcgggcccgt 2340  
agccaggctc gttgtgaacg tgtggaggag ctcaccttgg aggcaccgct ggtcctgccg 2400  
gagcggggcg ctgcggcggt gcagttggcg gtgggggctc cggatgaggc cggccggcgc 2460  
agtttgcagc tctattcccg aggcgtgat gaagacggcg actggcgggc gattgcctcc 2520  
gggctgttgg ccagggccaa tgcggtgccg ccggcggatt cgacggcatg gccgccggac 2580  
ggcgccgggc aggtcgatct ggcggagttc tacgagcgcc tcgccgagcg cggcttgacc 2640

tacggtcg	cg	tattccaagg	gctccgcgc	gc	atggcg	gc	acggcgacga	tatcttcgcc	2700					
gaattggccg	gg	tcaccaga	cgcctcg	gt	ttcg	gc	atcccgcgct	gctggacgct	2760					
gcactgcacg	cg	atggcgct	tggtgcttcg	ccc	gactcg	g	aagcgctct	gccgtttcc	2820					
tggcgtagcg	ccc	agctgta	ccgcgctgaa	gg	agcagcgc	ttc	gggtacg	gctctcgccg	2880					
ctgggctccg	gt	gcagctc	attgacgttg	gt	ggatgcca	cag	ggcgacg	agtcgctg	2940					
gtggaatcgc	tttc	gcgcg	accggtctcc	acc	gaccaga	tc	ggtgccg	tcg	ggcgat	3000				
caagagcggc	tg	ctgcacgt	cgagtgggta	agg	tcggctg	aat	ctgcggg	gatgtctctg	3060					
acctcctgcg	cg	gtgctg	tttgggcgaa	ccg	gagtggc	ac	gctgcgct	gaagaccact	3120					
gggtgtccaa	tc	gagtc	tcg	gg	ac	ctt	gcttcg	ttgg	cc	accgaggt	tg	ccaagcgg	3180	
ggttcagctc	ct	gg	tg	cg	gt	ccg	gc	gat	ccc	gagcgat	gc	aggagctg	3240	
ccgaccgccg	cg	gaagg	gc	ac	caacag	gc	gatggcg	tg	ctgcagca	at	ggcttgcc	3300		
gatgaccggt	tc	gtcagta	gcgcctgatc	ct	gctgacgc	at	cggg	cggt	ct	cgcgagtt	3360			
gctggagaag	ac	gtgctcga	cctgggtacac	gc	gcgcgctgt	gg	ggc	cttgg	cc	gcagcgcg	3420			
caagcggagc	acc	cg	gaccg	att	gc	ccttg	at	c	gat	atgg	acgacgagcg	ag	catcgcag	3480
acggcactcg	cc	gaagcgt	gactgcggga	ga	agcgcagc	tc	gc	ggtg	gc	gt	cgggagtt	3540		
gtgctggcgc	ccc	gcctcg	ccaggtgaag	gt	gagtgagg	gt	gaagcgtt	cag	gt	ggg	gat	3600		
gaaggcaccg	tg	ctgg	tcac	cg	gc	gaacc	gg	cggg	ctcg	gg	ccctgct	cg	cacgccat	3660
ctgggtcagcg	ccc	acggtgt	gcggcacctg	tt	gctcg	caa	gt	cgc	cg	gtg	gtg	gtg	gtg	3720
cccggagcgg	at	gagctggt	ggcgcgagctg	gag	caggccg	gc	gc	ccg	acgt	cg	cggtcg	3780		
gcgtgcgact	cg	gcagatcg	ggactcgctt	gc	gcggctg	tg	gc	gtcg	gt	gc	ctgc	gaa	3840	
aacccggttg	gg	gtggtg	gt	gc	acgcgc	gg	gt	gtg	ctg	at	gacggtgt	gct	gatgtcg	3900
atgtgcgcgg	ag	cgttgga	cgcggtgttg	cg	cccaaag	tg	gatgc	gc	gt	gt	tac	ctg	3960	
cacgagctga	ct	cgggaact	cggtctgtcg	gc	gttcgtgt	tg	ttctc	ctc	gt	tc	gcgggc	4020		
ctgttcggcg	gt	gcgggca	gagcaattac	gct	gcgcgca	ac	gctttc	ct	gg	atgc	cttg	4080		
gcgcattgcc	gg	caggccca	ggggctgccc	gc	gtgtgc	tg	gc	ctcg	g	ct	gtgggc	4140		
agtatcgatg	ga	atggcggg	cgacctcgct	gc	ggcagatg	tg	gagcggct	gt	cgcgggca	4200				
ggcattggcc	cg	cttcggc	accgggaggg	ct	ggccttg	tc	gacgctgc	cg	ttgg	ctcg	4260			
gacgaaccgt	tg	ctggcacc	ggtgcgactg	gat	gtcgaag	cact	gcgtgt	gc	aggcccga	4320				
tccgtgcaga	ccc	gattcc	ggaaatgctg	cat	ggcatgg	caat	ggggcc	aag	ccgccgc	4380				
actccgttca	ctt	ccaggtg	tgagccgttg	cac	gaacggc	tg	gcgggatt	gt	cggagggc	4440				
gaacgtcggc	ag	caagtgt	ccagcgcgtc	cg	cgcgata	tc	gcggtggt	act	ggggcac	4500				
ggcaggtcga	gc	gatgtgga	catcgagaag	cct	ttggccg	ag	ctgggttt	cg	actcgctg	4560				
acggccatcg	aact	ccgcaa	ccgtctcgct	acc	gccaccg	gact	gcggct	tccc	gcgacg	4620				
tggccttcg	acc	acggcac	tgcggcggca	ct	cgcgcagc	ac	gtgtgcgc	gc	agctaggc	4680				
accgcgaccg	cg	ccggcacc	gaggcgaacc	gac	gacaacg	ac	gccacgga	gccc	gtgagg	4740				
tcgctcttcc	aac	aggcgta	tgcggtggc	cg	gatacttg	ac	gggatgga	ttt	ggtgaag	4800				
gtcgtgccc	ag	ttgcgacc	ggtgttcggt	tc	gcctggcg	ag	ctggaatc	cct	gcgaaa	4860				
ccgctccagc	ttt	cccggtg	tcccgaagag	ctt	gccttg	tgt	gcatgcc	gg	cgtgatc	4920				
gggatgccgc	cc	gcacagca	gtacgcgcgg	at	gcgcgcg	ggt	tcgcga	tgt	gcgggac	4980				
gtttcggtga	tccc	gatgcc	tggattcatt	gc	gggagaa	cg	ctgcgcgtc	cg	ccatcgag	5040				
gtggcggttc	gg	acgcaggc	ggaggcggtg	ct	gcaggaat	tc	gcggggg	ct	cgttcgta	5100				
ctgggtcgggc	att	cctccg	gggctggctg	gc	gcacgagg	tag	ccggtga	gct	ggagcgt	5160				
cgcggggctcg	tccc	ggccg	ggtcgta	ct	ggacacct	ac	atccccg	tg	agatcacg	5220				
ccgaggttct	cc	gtggcgat	ggccaccg	ac	gtatgaga	ag	ctcgcgac	ttt	cacggac	5280				
atgcaggatg	tc	ggtatcac	cgcgatgggc	gg	gtacttcc	gg	atgttcac	cg	agtggact	5340				
ccgacgccga	tc	ggtgctcc	gacgctgttc	gt	gcggaccg	aag	attgcgt	cg	cagacct	5400				
gaagggcggc	cgt	ggacaga	tgactcctgg	cg	gccagggt	gg	actctcgc	gg	atgccacg	5460				
gtccaggtgc	cgg	gcgacca	cttctcgatg	at	ggacgagc	ac	gcgggtc	cac	gcacag	5520				

gcagtcgcga gttggcttga caaactcaac cagcgcaccg ctccggcaacg ctgacggggcg 5580  
tccttttagg accttctggg cggcaccggc caccocggcg gtgcccctt ccgtgggtcca 5640  
ggctcgccga tcttgacggc gcacgatgcg cggcaccgcg gctgatcgtg attccgctgc 5700  
cgctcgtggc catcggcctg gcgaatcatg tccttttcggg caacgtcaaa cgaattcgtc 5760  
cgagcccgca ttccgaggtg agggggcaccc ttgggtggct gagccgctca aggggtgccc 5820  
tcacctcgaa attcgtccga tttgggcggt ggacgcaacc ccggtggggcg tgggtcgtct 5880  
ttcttgttga cagagcgggt agaagccgct gacacacctg agaggaaaag gggagcatga 5940  
tgctcaagcg ccaccgtttg acgaccgcca tcaccggcct tctgggggga gtactgctgg 6000  
tcagcggctg cggaaaccgc gccgcacttc agtcctcgcc ggcgcccggg catgacgcgc 6060  
gcaatgttgg tatggcctcg ggcggggggcg gcggggacat cggcacgtcg aactgctcgg 6120  
aggccgattt cctcgccacc gcgacaccgg tgaaaggcga ccccggcagt ttcacgtggtg 6180  
cgtacgggaa ccggtcggac aagacctgca cgatcaacgg cggcgtgccg aacctcaagg 6240  
gcgtggacat gagcaactcg ccgatcgagg acctgccggt cgaggacgtg cggcttccc 6300  
acgcgcccaa ggaattcacc ctccagcccg gtcagagcgc gtacgcggc attggcatgg 6360  
tcctggccga cagcggcgac ccgaacgcc atgtcctcac cgggttccag tcctcgtctg 6420  
cggacatgtc cgaggccag ccggtcaacg ttctcggcga cggcaacgtg aagttcgccg 6480  
cgaagtacct gcgagtcagc tcgctgggtg ctaccgcaga cgagctgcgc taaaacccat -6540  
gtgagtcctg cagattcgac ctgcctcgtg ggcgcctccg gcgaagcgtc cgtacgtttg 6600  
tcgttgtgac cagcgttgtt cacgtccggg cgcagcgtg gtacatactc aggcgtctcg 6660  
ggcgccctca acggggcctg gcatccgggg ccgtcagagt cggcggcgct gacgcgttct 6720  
ctgtcggggc ttgtcacgcc gccggcctcg aaccggctcc gcccgctcg agccggtggt 6780  
ccagcgcggt gtggcggcgg ccggagccga cggtgccac cgcctgccc agggcctttt 6840  
tcgaaccgac gaggaccacg acctctctgg ccgggtgac cgcctgtag agcaggttgc 6900  
gctgcagcat catccaggcg cttgtggtca aggggatcac caccgacggg tattcgcttc 6960  
cctgcgaacg atggatggtc accgcgtagg cgtggaccag ttcgtcagat tctgtgaagt 7020  
cgtagtcgat gtctcgtcc tcgtcgttc gcacggctcat ggtctgtgct tcgttgtcga 7080  
ggcgggacac gacgccctgc gtgcccgtga acacgccgtt ggcgcccttg tcgtagtgtg 7140  
tgcggatctg cgtgaccttg tcgccgacgc ggaagatccg tccgccgaac cgcctctctg 7200  
gcaggccctc cctggccggg gtgatcgctt cctgcaacag ctggttcagc gcgcctgcac 7260  
ctgcggggcc tcgatgcac ggggcgagga cctgcacgtc ggtgcgcggg ttgaaccgga 7320  
acttcgcgg aatccggcg gcgacgacgt cgacggtgag ctccggcggtc ggttcgcttt 7380  
cctctacgtg gaacaggaag aagtcggtca gcccggtgtg cagcggatag tccccggcg 7440  
tgattcggtg cgcgttggtc accaccccg actcggcggc ctgccggaac acctcgttga 7500  
gccgcacgtg tggaatcgg gtgccagggg cgagcagatc gcgcagtacc tcaccggctc 7560  
cgaccgacgg gagctggctg acgtcgccga ccagcagcag gtgcgcgccg ggcgcgatcg 7620  
ccttgggccag tttgttggt aacagcaggt cgagcatgga cgcctcgtcg accacgacga 7680  
ggtcggcgct cagcgggttg tcccggtcgt aggcggcgct cccgcccggc tggagtgtga 7740  
gcaggcggtg caggtcgcgc gcgtcgtgct cgggtgagctc ggtcagccgc ttcgcgcctc 7800  
gtcccgctcg cgcggcgagg atcaccttg cctttttcgc ctgagctaat gcgatgatcg 7860  
accgcacggt gaagctcttg ccgcagcctg gacctcgggt gagcaccggc accttctcgg 7920  
tcaggggccag cttgacggcg cgtcctcgc cctcggcgag ttcggcaccg gtagcgcggc 7980  
gcaaccagtc gagggccttg tgccaatcga cgtcggcgaa gacgggcac cggtcgcgcg 8040  
tggtgttcag cagccgggac agctggttgg ccaggcgac ttcggcgcgg tggaagggca 8100  
cgaggtagat cgcgaccgtc ggcacctcgt cgtcatcggt ggggatctcc tcgcggacca 8160  
caccttctc ggtgacgagt tcggcgaggc attcgatcac cagcccggtg tcgacggcga 8220  
ggatcttcac cgcctcggcg atcagctcgt tctccggcag gtagcagttg ccgtcgcggg 8280  
tggaactcga cagcgtgaac tgaaggcccg cctttacccg ctgcggggag tcgtgcggga 8340  
ttcccaccgc tttggcgatg gtgtcggcg tcttgaaacc gattccccc acgtcgcctg 8400

ccagccggtg tggctcttcc ttgacgggtc ggatcgcgtc gtcgtggtac tgcttgtaga 8460  
tcttcaccgc cagcgaggtc gagacgccga cgccttgacg gaagatcatc acctccttga 8520  
tcgccttctg ctccctccac gcgtcggcga tcagcttcgt ccgcttcggg ccgagcttgg 8580  
ggacctcgat cagccgcgcg ggttcctgct cgatgacgtc gagcgcggcg acgccgaagt 8640  
ggtcgacgat cttctcggcg agtttggggc cgatgccctt gatcaggcca gaccccaggt 8700  
agcggcggat accttgacg gtcgcaggca gcacggtcgt gtagtcgtcg acgtggaact 8760  
gccgcccgtg ctgggggtgc gacccccacc ggccgcgcgt gcgcaacgcc tcgcccgggt 8820  
gcgcgcccag cagcgcgcgc acgaccgtca ccaggtcacc gccccggccg gtgtcgatcc 8880  
gcgcgacggt gtagccgctc tctcgtttgg cgaacgtgat ccgctccagc gtgccctcca 8940  
gcaccgcagt ccacgtggcc gactcccgtc ctttttccac cgacaacacg tatcacgaac 9000  
ggctgtcaag caaacggcg gtcaccacat gcagcggcat ctcccgaacg cctcggggtc 9060  
cggcgtcagc ggggtggcgt tcgcgatgcc ttggtgcggc cgggtgggagt tgtagatttt 9120  
ttcgtcctcg cgcagggcct ggagtaggtg ccgctgggtc cagatcaggg gaaagcgccg 9180  
ggataaaccg gcttgacgga ggagatggaa gagctctacg tcgaatggcc agcgaccac 9240  
gacgaccccg agtcatgctg cgacgacctg cgaggcgtg gcgaagcgtt gacagggggc 9300  
acgtgggggc gtgcgatcag gtcgcggcgc cggctcagcg ccgtgtcccg acgcaccagc 9360  
agccgcgacg gactggttcc gctcctcgag cgcggtgtcc tcgctcgtct ttttcgagct- 9420  
cgatgcgggc cgcgttcgcg aggtggccag tccagtcggt cggcagtcgt ggtcgagtc 9480  
tgccctttga acatctggtc aattcagctt tgaatgatcg ttccggcaggt attgtcctgc 9540  
ccgccaattg tttgctcgcc gtgcgcctcg ggacaggagc gcggcggttcg gggacggcgt 9600  
gctggcgaag gtgtcgattg gtcgggacgg ccggaccagg atcgccgacg tcgtgatcgc 9660  
cacgagcgtc ggcgatgcc ggtactatct ggagcaccag cgagtggaca acgacttcca 9720  
cgggcgcggt gcgctcctga tcatgaatca agaactgaga ccgtcagggt gaacgtgccg 9780  
tcgtggggca ctgtgaggat gtgcatacaca cggccgggggt aggtgatcga ggaggccctg 9840  
atcgccgggt gccgcggctg tgcccagtcg aatcgacgca ccacctggcc gcgggtgagc 9900  
gtcagccctt gcgcgtcgag tgcgactttg tcggagtcce tcaccacca cgggatctgc 9960  
tcggttgcgg gtgcgttcgc ctgcaccgac ctggtgatcc ggtcggggccg caccgtcaca 10020  
tcggtgacga ccgagccatt cgcggtgcgg tagcgggaacc ggtatggctc gccgcccagg 10080  
aactccgctt gctgcgggtc gttcgcgtcc ggcacctggc ccgggaagac cgtcgcccag 10140  
caggcgtggt cgttggttgc catcgactgg acgaccgtgc ccgcgaccgg gtgccacagg 10200  
aacgtcagac cggtgccggc cagcgaggtc gcgcgcctcg tgccgaagta gccgcccagg 10260  
tagaactgcg gcctgcgtag gaacaggaag tcttgccctt ggtcgtgcg gatctcgggtg 10320  
aagtgcgctt gttggtatgg aagtttcgcg atcgcgccg tcgctgggc gcgggtcgga 10380  
tagcgtcgc cgtagatcgc gtgggtgagg atgcgcgtg aggtgtcctg tttgacgagt 10440  
gcgggcaccg gtgcgggtc ctcgcccgag gccgctcgt cggcggcgag gtcttcgcgg 10500  
ctggacagga acgcgcccac gttcggcacc actggcacga actggttgtt cagcgcggtg 10560  
cgttcgggtt ccggtcgac gtcggcgtag aaccgcgcgg aggtgcggga gttggggcg 10620  
acgttggcga accagccgga gccgtcgggt tcgcgcagca ggttgtagct gagccagtcg 10680  
gtgtacttcc gcgccatcgt gacgaggttg gcgttggtg actggcgcca cgcgtcggcc 10740  
atctccgga gcatacctc gaagttgtag ttggtgtccg cgcgggtcaa ctcgtagaag 10800  
tagccgaag gactctgcc gttcgcggca aggcgcgcga acgcctcgggt caactgctgc 10860  
cgcagcgcgg ggtcggggtc ctggttcagc gcgagcgagg agccggcgag cccggcggtg 10920  
acctggttcg cgtagtggat tgtgggtgc caggcgctac cgttgcccgg gttgaggagc 10980  
cacgtcatcg ctttgcgag cgcgaggtg atctgcgact ggcgttgcg caggatgttc 11040  
gccgcgcga gcaggcgtt ggtcttgctc aggtaccca acccgaagcc ggtggcgcg 11100  
agcccggtc cggtcggcga atactccggc cagagccgt catcgtgctg caaccctagg 11160  
tagtgtcga gcccgcgctc gagcgcggcg agcagagtcg cgtcgccccg gtacgggttc 11220  
cacgtgcggg attgcgtggc gaaccaggcg agcgtgtaga cgtgctcctg gacgcggcg 11280

ttgtaggaca cgcgcggtg ggcgcaccag cctccggcga agaagcgtc cgagtccatg 11340  
 tcggcgacca tcggcgagac cgcggtgagg tacgacgcga atcgctgctc ctccgggtgcg 11400  
 aacaggcgcc ggttcggtgt ggtggcgccg ggggcgggca gtgcgagtgc gcgggtcggc 11460  
 agtgcggcga gcagtccgag agcggcgccg ccggtcatca ggctgcgacg gctgaacgta 11520  
 gtcacgggcc tacctccttg tggccgatca accctcaccg gctgcgtagc cgcacgtcaa 11580  
 gatgataatt cgaattatta tgggcttgac gacgcgtagg ccgacgacgc agaattcctg 11640  
 ccaattcgta ttggcaagcg ggggtgctcg tggcccgacg gctcaccacg caggacatcg 11700  
 cccggatggc aggagtcagc caggccacgg tgtcgctggg gctcaacaac cggaaggacg 11760  
 gcaacgtccg gatcgcgccg gagaccctg cgcaggtact ggaggtgatt cggaagaccg 11820  
 gctacgtcgc gaacccgatc gcccgagga tgcgcgatcg gcacaaccgc atcctcggcg 11880  
 tgttcaccta cgaggcggtg ttcccgagca ccacgcgaa cttctaccag tcgttcctcg 11940  
 aaggcatcga ggaacaggcc gaggaggtcg gctgtgacct gttgtgttc accagcgcca 12000  
 aggccaccgg gagcgccg cggattttcg gcgacgacg ccgggtgcgg ctcccgacg 12060  
 gcagtctgct gctcggtcgc acggtcgacc gcgacgacct gaccagctg ctcccgaaa 12120  
 gcatcccgta cgtctccatc ggacgacgcg acgacgcggg cggtcgggtg ccgacgctg 12180  
 gggcgacta ccgaccgcg gtgcgagacc tggtgaccg cgcggtcgcg ctccgccacc 12240  
 gcgggttcgc gtacgtggg tctggtggg gcgcggagtc gtcccgggat cggctgcgag 12300  
 gcttcgcga agccgttgcc gcacatggc tccaaggat gcatgtggag accccacagc 12360  
 tcgatcagct gcgcgaagcg ggcgtcaccg ctgtgctcac cgaagaggtg tcggacggg 12420  
 ccgcgctcgt gctcgccggg cgcgaacgcg ggctctccgt gccgggcgac ctcccgctg 12480  
 tctcgctcgg tgcgctacc cgtcgccac cggacgacga cgtggcgctc accggtttcc 12540  
 gcatccccag gcgcgagatg gggcgccggg cgggtcgagg gctgaccgag gtgctcgaaa 12600  
 acggcaccac accgcaagaa ctgctccgt gcgagttcgt cgagggctcg acgctggcg 12660  
 caccacgct ttgaccagga ggaactgttg ctcgaccaca ccacggacgt tgcgtcgtt 12720  
 ggcgcgccgac tcggcggtgt cgcgcgcga ctccggttc tgcgcgcggg ccgcgggtc 12780  
 gtgctcaccg aggagtacga ctggctcggc ggccagctga ccagccaggc cgtgccgcc 12840  
 gacgagcaca gctgggtgga gcgcttcggc gtcaccgcga gctaccggg gttacgcgac 12900  
 ggcacccg actactacc cgcgcactac ccgctgacct cgcgcgcacg ggcgtggcg 12960  
 gagctcaacc ccggtgcgg caacgtgagc aggtctgctc acgagcccc cgtcccgctc 13020  
 gcggtgatcg acgagatgct ggcgcgttc cgcggcagtg gcaggctgac cgtgctgcag 13080  
 ccgtaccggc cgggtggcgc gcacaacgac ggcgaccgga tcgtgctcgt gaccgttgcg 13140  
 caccgcgaca ccggtgaaca gatcgagctc tccgcgccgt acatcctgga cgcgacggag 13200  
 acgggtgaac tgcttcgtt gtccagcacg gactacgtca ccggcttcga gtccactctg 13260  
 gcacccggcg agccgagtgc gcccgacgtc gcgcagccg cgaacatgca ggcggtgctg 13320  
 gtgtgcttcg tggtcgacca cgtcgacggc gaccacacca tcgacaaacc ggcgcggtac 13380  
 gacttctggc gcgcgtacca gccggaactc tggggcgacc ggatgctgtc gttccgctcc 13440  
 cccaaccgc gcacgctcgc gatctccgaa cgtacgttca cccgaaccc ggacgacgac 13500  
 ccgctcggcg tcgtgtcggc ccagcggtc agtgccggtg acagcaatct gtggacgttc 13560  
 cggcgcatcg ccgcgctcg caacttcgtc gagggtgctt acgacagcga catctgctg 13620  
 gtgaactggc cgatcatcga ctacttcgag tcgccggtga tcgacgtgcc ggacgcgcgac 13680  
 gcgcacatcg ccgcggcgcg ggaactctca cgttcggtgc tctactggct acagaccgag 13740  
 gcgcgcgcg cagacggcg caccggctc cccggcctcc gctgcgcgg cgacgtcac 13800  
 ggcagcgcg acggtctcgc gcaggcgccg tacatcccg agtccaggcg catcagggcc 13860  
 gagcacacga tcgtcgaaca ggacctctc ctcccgctgc gcggcgacaa ggggtgcggtg 13920  
 cagcacgcg acgcggtggg tgtcggcatg taccgtatcg acctgcacc ctccaccggt 13980  
 ggcgacaact acatcgacgt cgcgagctgc ccgttcgaga tcccgctcgg cgcgctgatc 14040  
 ccgcaacggg tggagaacct gctaccgcg ggcaagaaca tcggcaccac ccacatcacc 14100  
 aacggttccc accggtcgca ccagtcgag tggaacgtcg gcgaggtcgc gggcgcgctc 14160

gctgccttct gcctggcgca ccgagtcacc cctcgcgcg tgcgcaatac ccctggcctg 14220  
 ctcgcggaact tccagcagtg tctggaacgc gacggggctg agctccgctg gccggacgtg 14280  
 tccggctact gacgcaggga gacgaaaatg acaaagctgt cacgacgact cacggcactc 14340  
 atcgctgcag ggctgttcgc cctcaccggc tgcggtgggt gatcaaccgc acagtccgga 14400  
 ccgaagtgcg tgcgcatgac cgtgtggact gccaacgcgg cgcattctca gctgctcaac 14460  
 gacatcgccg ccgagtacaa ggcctcgac cgggacatcg ccgagatcaa gttcgactcg 14520  
 gtgcccgcgg acggctacac caccacgctc accaccaga tgcgcggcgg taacgcgcgg 14580  
 gacctggcct ggatcctgga ggagtcggca ccggacttcg tggcgctcgg tgcgctcgcc 14640  
 ccggtgcgcg gcaagatcga gaaggccgac gagctcgtgc cgtccgcgac gaagctgtgg 14700  
 gagaaggacg gcgaactgta cgcctaccgc ttctccacct cgcggttcgg cttgttcgtc 14760  
 aacaccgacc tggatgaagg cgcctcggcg gactggacct gggaccaggc gatcgcggtc 14820  
 gcctctgctg cggcgccgc ctccggcaag ggcggcctgg tactgccgga cttcaagtac 14880  
 cagaactggg cagtgtgtgc ctctatctgg cgcggtctgg gagctgatgc gtggagcgcg 14940  
 gacggtcgct cgtgcgggtt ctccagcagc gagatgaac acgcgatgtc cttcttgac 15000  
 aaggccatct taccgacaa ggcgattccg ggcggcgga cgacggtgga cttcttcgcc 15060  
 ggcgacgcgg cgatggcgat cggccagatc tcccggtcca gtgcgttgaa ggacgcgaag 15120  
 ttcggctgga cgtgtgtgc gctgcggcc ggtccgaagg gtgactacgc ggtgatcggg 15180  
 caggccggga tcggtgtgt gaagcagtc cacaacgtc acgcgcgac ggacttcctc 15240  
 gccttctca ccaaccagac caactccgc aagctcgcgc agttcttccc cccggcgcg 15300  
 tcgtcgctgc tcaacgcgga gacgtcgc aagagcaacc cgttgatcaa ggccgagcaa 15360  
 ctgcagtcgg tcgtcgctga cggcatcaac aaggcgctcg tgaagccgag ccacaaggg 15420  
 caggaggagc tgaaccagac gatccgcgc gcgctcgacc cgttggtgaa gccggacgcg 15480  
 aacgtgcaga acgtgtgaa cgacgtgtgc accaagatca aaccgctgct ggagaacaag 15540  
 tgacggcggt cgcacactcc acgcgggccc tagcccgcaa ggggacgtc tactggacgc 15600  
 agcgcaggcg cgacaatctg gttggttacc tgttcgtgc gccgcgctg ctcggcagca 15660  
 tcgcttcgt gctggtgcg ctggccttg tcggtggtg cagcctcaac gagtggaaag 15720  
 tgctcgccg caggttcgag ttcgtcggcg cgcagaacta ccaacagctg ctcgccgacg 15780  
 agaagctgcg cgactcgctg gtggcgacca cttggttcgc cgcggcctg gtcgtgctca 15840  
 acctgtcact agcgtgctg ctggccgtgc tgetcaacca gaagctgagc ggcaccacgg 15900  
 tcttcgcac gttgttctc tctccgctc tcgtgtcgt ggtggcgtg acgtggtgt 15960  
 ggcagctgat actgcagcc gagggcagcg tcaacggact gctcgggttc ttcggtgccg 16020  
 acgggcccga ctggtgcgc ggtgagtcga ccgcgatggt ctcggtgatc gtcgtgcagg 16080  
 tgctcaagaa cgtcggcctg aacatggtgc tgttcctgc ggcgctgcaa ggcgtgccgc 16140  
 agccgttgta cgaggcggcg aagatcgacg ggcggggtgc ctggaccggg ttcgcgcga 16200  
 tcacctgccc gttgatcagc ccgacgatcc tgctcacgtc gatcatcacc atcgtcggct 16260  
 cgctgcaggt gttcgcgcag atcgcggtgc tcacgcagg cggtcggggc acgtccacga 16320  
 ccgtactgat ctactacctg taccagcagg cgttcagtt ccaccacttc ggctacggcg 16380  
 cgacgatctc ggtactgctg ttcgtcatcg tcgccgact caccctgctg cagtggcaga 16440  
 tgcgcgggaa gtgggtgctg catgaggctt agggtaaga tcacgctata cgggctgctg 16500  
 tgctgctgt gcgtgcctt cgtgttccca acctggtgga tgatcacctc gtcgatgaag 16560  
 cccatcagcg agatcttct cactcgcgg ctgcccgcg agtggacgtt ctccacctac 16620  
 cggcaggtgt tcgagatgca gccgttcgcg cagcagttact ggaacagcct ctacatcgcg 16680  
 gtgatcgta cgacaggcac aatggcggtg gcggcgatgg ccggatacgc gttcgcgcgt 16740  
 atccggttcc cgggccagaa cgtgtgttc gtggtcgtc tgatcggcct gctcatccc 16800  
 agcgaggtca ccatcgtgc gctgttcaag atgttcagc cgtcggcct gaccaacacg 16860  
 cactggccgc taatcatcgt gccgatcttc ggggcgcct gtgtgctggc gatcttcac 16920  
 atgcgccagt tcttcacgc cctgcgctg gagctggagg aggcgcgcg gatggacggc 16980  
 ctcggacgag cgggcatctt ctaccgggtg gccctgccgc tgtccgggc cgcgctgggc 17040



gccgtgcga tcttcacgct cctgcactcg tggaaacctct acctggagcc gatcgtctat 17100  
 ctgtccacac cggacatgta cacgttgccg caagcgctca cgcagttcgt cgacaactac 17160  
 ggcgggccga tgtggaacgt gcagctcgcc gccgcgacca tgaccgcagt gccgattctc 17220  
 gtcgtgttcg tgctggcgca gcggcagttc atcgagggac tcgcgcacac cggggttaag 17280  
 gggtagtcc cgcgatcgac ccggagatcg ccgcgctggt agacctcgcc agaccaccgg 17340  
 tgacgccagc gggaaatcgac gcggtccggg cgggaggacg cgtcgtcacc gatgccgagc 17400  
 tgacgcgcgg tggcaccgtg acgttcgccc atgccgatgc cgatggcgtg cccgtgcttg 17460  
 tgctgcgccc agcgggtgtg ccacgtcttc cggtagtga cctgcacggc ggcgggatgg 17520  
 tcgcgggcac ccggcgacac gacctgcacg tgctggccga gtgggtgtcg gagctgggtg 17580  
 tcgtgctggt gtgcgccgag taccggtcgc caccgagca cccgcacccg gcgcctgcgc 17640  
 aggactgctt cggggtgctg gagtggtatg cccgcaacgg ctctggcccg cctgtcgttg 17700  
 cgggtacatc ggcgggcggc gggctggcgg ctgcggtgac gctgatggcc cgcgacctcg 17760  
 gcggtccgcc gatcttgccg tcggcgcgca actgctcgaa cgcgtccagc gtctccgcca 17820  
 acggcgtgct accgcgccag tgcagcaggt agacgtccag gtgatcggtg cccaaccgcc 17880  
 gcagactgcg ttgcacgccc tcgacggcgc cagccgacc cgcgttggtg ggatacacct 17940  
 tgctgaccag gaaaacctcg tcgcggcgcc cggcgatcgc ggcaccgacg acctcctcgg 18000  
 cgccccgct gccgtacatc tcggcggtgt cgatcagccc caggccgagg tccaggccgc 18060  
 gccgcagcgc ggcgacctcc gccgcgcgt gccacgccc ctccccatg cccaggtgc 18120  
 cctgccccag caccggcaat tcctgcccgc cagcgagttg aaggcctcgc atgcacactc 18180  
 cgtttcagtc agcgtgctg cggcagcacg gccaccagct cgaagccacc gtctcggac 18240  
 ggccccggcg ccaaccgccc accggccatc tgcaccgct ccgacaggcc ggccagcccg 18300  
 gaccgcgcgc tgccgacctc gatggcgggt cccgcgcggt cggggggcgc gtttgtagc 18360  
 gcgacgtgca gcgagccctc gcggaacacc agccgcacct tgatgtcagc accgggtgcg 18420  
 tgcttggaag cgttcgtcaa cgcctcctgg acgaccgggt aagccgcgcg ctcaaccgca 18480  
 ggcgacaccc cgtgcgggtc gcggtatctc gaggaaagct ccacatccac accggcgcg 18540  
 cgcgcgccga cgaccagctc cgggatcgtg ctcaaccggg cccgctgcct gggggcgccg 18600  
 agcccgagcg tcgcgcgcat ctgctccagc gcctgcttcg ccaaccgcgc cagccgcccg 18660  
 gcggtctcct tcgcaccggg atcctcgggt gtggcgcca acgcgcgca ctgcaccgcg 18720  
 attaaggtca cgtgatgtcc caccgcgtcg tggatctccc gcgcgatccg ggcccgctcc 18780  
 tcgcgcgcgc ccgtctcggc ctgcgctgc aactcgcct cggtagccgc ctccagcttc 18840  
 cgcgactgt ccgccagctg ctcccgcaac gccaccagcg cccccaatgc cgtcggcgcg 18900  
 cctgcccga gcagcacgaa cgccacggtc aggatgatcg acccgacccc cagcgactgc 18960  
 atgatcatca ccggtgccac ggcgaccgcc gcggtcagca acaccagcc gaccaacatc 19020  
 tggatccgcg gctgcgcgcg ccccaacgcg tacaacgcca cgaccgtcg cgcccaaccc 19080  
 aaaccaccgg ccaacgctgg catgcacagc agcaccgca tccgcggcca ccgccgcaac 19140  
 ggcagcagcg ccgaagccac caaaccggac accgtcgcct agccgaacgg gtcgggcgcg 19200  
 gccagcacag ctgcgcgcgg aatcccgcgc accagcgcct caacgcgcg ccgagaccag 19260  
 cgcaccgaac tcaccggaag ccgcgcgcca cgcgctgcgc gatcaacgcg gcctgcaccc 19320  
 ggttctgcgc gccgagcttg ctcaacaccg tcgacacgta gctcttgacc gtgcctcgg 19380  
 tcagcccgag ccggaacccg atatcggcgt tggaccgacc atccgccaac aactggagca 19440  
 cctgcgcgtc ccgcccggac aacccctgca ccagctgcgt ctgctgcgcg ttgtcggaca 19500  
 gactgcggaa cctgggcaac aacgcgcgag tgatccgcgg atccagcacc gcccaccgg 19560  
 ccgccagatc gtgcaccgcc cgaaccagca ccgcgggctc cgcgtccttg agcaggaacc 19620  
 cgtgcgcccc gaaccgcaag gcacggcca cgtaatcgte caggtcgaac gtcgtcagca 19680  
 ccgcgggcac cggcggtatc gccaacgtcg ccaagtcgag caacgcgcgt agcccgctc 19740  
 tgccgggcat ctgcacgtcg atgaggcgca cgtcgacgcg gtacgcgcgc accgcggcca 19800  
 gcaactcatc cccgttgccc gcctcggcga ccaccggat cgagccgtcg cctccaaca 19860  
 acaccgcag gccacgcgc agcatcgct cgtcatcggc aagcaaacg cggatcggcc 19920



ccccgccgtc gtcagcccgc acctctcggg cacctccctg ttccgagggg cgccagcaga 19980  
ctagcggggc cgcgaaagcg cagcgcacac gccacccgag cgcactgtgt cagcggcgcc 20040  
cgcccttggc ctcccgctcc cgcacccgca ccgagatccg caccgggcta cccacgaacc 20100  
cgaacgtctc ccggaacctg cgtccagga accgccggta cccggcctcc aagaacctcg 20160  
tggtgaacag cagcatcgtc ggcgcccgcg actgcgcctg ggtggcgaac aggatcttcg 20220  
gctgcttgcc gccccgcacc ggcgccgggt tggccgcgac cagctcggaa agccaacctg 20280  
tcaaccggcc ggtggagatc cgggtgtccc acgaatccag cgccgtgcgc agcgtcgggg 20340  
ccagtttgcg caccgcgcga ccggtcttcg ccgacacgtt gaccgcgtcg gccacccgca 20400  
cccgaccag ctgcgggtcg atctccttct ccaactgggt gcggcggtcc tcgtcgacca 20460  
ggtcccactt gttgtaggag atcaccaatg cccgaccggc ctgcacgacc atagtgatca 20520  
cccgcaggte ctgctcgctc aacggctcgg agcgtcgat gagcacgac accacctcgg 20580  
ccgctcgat cgcgcccttg gtgcgcagcg acgctagta ctcggtgccg cttgcggtct 20640  
tcacgcgctt gcgcagcccg gcggtgtcga cgaaccgcca cactggccg tccagctcca 20700  
ccagcgagtc caccgggtcc accgtggtgc ccgcgacgtc gtgcaccacc gaccgtcttt 20760  
ccccggtcag cttgttcagc aggtcgcact tgcccacgtt cggcttgccg accaatgcca 20820  
cccgcgcgg cccaccggte gcgccgaaga tctgcgcgg cgtctccggg aacacctcca 20880  
ggacggcgte cagcaggtca ccggaaccgc gcccggtgaa cgcgctgacc ggcacgggt 20940  
cgcccaaccc cagcgaccac agcgaatgca cgtcggagac gcctcgctgg tcgtcgacct 21000  
tggttggcgg cagcagcacc ggacgcttgg accgccgcag caccctggcc acggcttctt 21060  
cggctcgggt ggcaccgacc cgggcgtcca ccacgagcag caccgcatcg gcggtgtgca 21120  
tgccacgttc cgcctgcgcg gccaccgacg cctgcagtcc ctgcgcgtcc ggctcccagc 21180  
caccgggtgc gaccaccgtg aagcggcgcc cgttcacaaa cgcgtcgtag gccaccgat 21240  
cccgggtgac accaggggtg tctgcacga cggcttcccg acggcccagc agccggttca 21300  
ccagcgtcga cttgccacg ttgggacggc cgaccacggc cagcaccggc tgcggcttcg 21360  
tcggctcctc gccctcgacc gcttcgacct cgtcgaactc agcccactca gcctcgctcg 21420  
accacgtgcc gtccaggcct tccaccgact cgtcggtcac gctcgtttcc ctctcgaata 21480  
gtcttggctc tggctccgca ccagccgatc cagctcggcg atcagctcga ccagttcgg 21540  
gcgcaccgcc tcggctcgca cgaccagccc ggcgccgctt ttgcccggcg gcagcacgat 21600  
cggctccccg aacaacacgt cgtacctcgg tagcagccta cggcccgggc catccgggcg 21660  
gcgcgtgccc cggcaggcca ccggcaggat cagcgctcc gacgtccgcg ccaaccacgc 21720  
cgcaccgtgc tgcgcattgg tcacatcgcc atcaccccg gtgcctccg ggaacacccc 21780  
gacgagacca ccggcccgcg gcaccgcgac cgccgcccag agcggggcgc ggtccggcgc 21840  
accgcggcgc accccgatct gcccgatgcy gcgcaggaa cagcccagcg caccgcgaa 21900  
atctcctgc ttgatcagga acaccgcgtg gcgcgggatc attccgaaca gcagcggccc 21960  
gtccatcatc gaactgtggt tggccaccag taccaccggg cccgtgcgcg ggatccgctc 22020  
cgcgcctggt atccgcaccc ggtacggcag ctgcacgatc cggcgcgaaa tccactggcc 22080  
gcaccggtgc attccaccgg acgcgcctc gggcagggtc tcgtcgtca tggcgctcgc 22140  
cgctccgccc cgaccagcag gccacggcac tccacgtgct tgcgcaactg cgtgagcacc 22200  
tccaccacac ccaggtccgt ggtgtcgagc tccaccgct cgtcggcctt gcgcagcggc 22260  
gccaccttc gccccgagtc cagcgcacg cggcgctgca cgtccgcgtg ggtccggctc 22320  
aggtcacccg cccgtccctc ggagacgtc tgcgcgtgc ggcgtgcgc gggggcgtgc 22380  
gcggacgcgc tcaggtagac cttcaaacc gccatccggg ccaccacggt gccgatgtcg 22440  
cgtccctoga cgaccatgcc gcctggcgac accagcgcct caccgatcag ctcccgctgg 22500  
tgggcgacca actgctccc cactcggcg accgccgaca ccgccgacac cgcgcgggtc 22560  
acctccggcc cgcggatctc ccgcccacg tccgagccgt ccaggaagat ctcagggcc 22620  
tccggtatcg tgcaccacgt cagccgcgt gcgcgcacca cgctgccac tgcggccgga 22680  
tccgacgggt cgacctcgg gcgcagcacc gccagcgtca cggcgcggt catcgctccg 22740  
gtgtcgaggt aggtagcccc caatgcgaac gccaaactgc gggacaccgt cgacttgccg 22800

gtgcccgaag ggccgtcgg cgccaccaca ccaaggagcc tggcgtcgc cacgtgcatg 22860  
 tctttcccgga gccggaaccg atcggcccga cacccggccg gaccaaccgt ccattctgcc 22920  
 tggtcgctc ctgccgatcc caatcgcatc cagatccaat ccggactgcg ggacaaatac 22980  
 gaaagttcta cgaataaggt cttgggtgtg gatgtaactg tgaccccctt gccagggctc 23040  
 ggcacgcagc aggacttcac caccgatcc ggccaccgca tcggcgtgat cacctaccgg 23100  
 gacggccgct tcgagctgat cgtctccgac cacgaagacc cagacaaggt ggccgctcgc 23160  
 gtcgccttga ccacctccga gaccagcacg ctggccaacc tgctgggcgc cccgcagctg 23220  
 gtcgcccagc tcaacgagca gcacagcgag gtcgcccggga tcaccacctg gcagctgtcc 23280  
 gtcgccccgg gctcccccta cgccggccga accctcggcg acacggagat gcgcaccgcg 23340  
 acctccgtgt cgatcgtcgc ggtcgtccgg gacggcaccg tgaccccctc acccgggccg 23400  
 gacttccagt tctccgcccg cgacctggtg gtggtcgtgg gcaccgcgaa ggggctacgc 23460  
 gcagccagcg aaatcctgga aaagggtga tccggcccgt cccgaacact tccccggcca 23520  
 cctgaccggg acacgcgtca ccaacggcgg tccagcgcgt gccgatcccc taacaactcc 23580  
 atacgcaggg cacacgcgag aggaaacacg ttgcaagaca cggcgatctc cctaatagaa 23640  
 ctgggtgcgg ttttcttcgg tctgggcata ctcgggagac tcgcatggaa aatcgggggtg 23700  
 tcaccgatcc cgctctacct gatcggcggg ctggccttcg gcaccggtgg cctgggtccc 23760  
 ctgcacggca tcgaaccgtt caccacctc gcctccgaga tcggcgtggg cctgctgttg 23820  
 ttgctgttag gcttgaata ctccgcggt gaactcgtca ccggcttacg ccgtcctgg 23880  
 ctggccggtg tcatagacat cgtgctgaac gcggtacccg gtgcgctggg ggccctgctg 23940  
 ctgggttggg gtccgctcgg cgcttcaca atggccggcg tcacctacat ctcttccctc 24000  
 ggcacatcgc cgaaggtgct cggagacctg ggtcggctcg gcaaccggga aacgcgggtg 24060  
 atcctgtcga tctgtgtctt cgaagacctg gcgatggccg tctacctgcc gatcctgacc 24120  
 gccgtgctgg ccggggtcag ctctctcggc ggactgaccg cggtcggggg gtcgctggtc 24180  
 gtgatcacc tcgggtcgtg ggtcgcgctg cgcttcggca agtacgtgtc cgcgctggtc 24240  
 gacagccccg acccgagggt ctctctgctg cgcctgctgg gggcggcgct gctcgtcgcg 24300  
 gggatgcct cggaetgca ggtctcggcg gcggtcggct cgttctgct gggcatcgcg 24360  
 atctccgggt cgaccgcggc gaacgcgacc cgcgtgctgg agccgctgcg cgacctgttc 24420  
 gccgcgctgt tcttcgtcgt gttcggcctc aacaccgatc cgagccagat cccgcgggtg 24480  
 ctgcccgttg cgctgctgct ggcggtggcg acctgcgcga cgaaagtgtc caccggtgg 24540  
 ttcgcgcgcc ggatgcaggg tgcggccgg atgggtcggc tgcgggcccg tgccgcgctg 24600  
 gtggcgcgcg gcgaattctc catcgtgatc gccgggctgg ccgtcgctc cggcgcggtg 24660  
 ccgggcgagc tcgcccct cgccaccgcc tacgtgctgc tgatggcgat cgtcggcccg 24720  
 gtcgcggccc gcgtcgtcga gccggtggcc cggatgttcc tcccagcca ggcgaagaag 24780  
 gcctgaccgg aaaagcctcc ggctgtccc agcggacggg ccggaggctt tctcaacgct 24840  
 ggcgataacg cgggcgggtc tggcaaacgc aacaccggtg cgcgaacagc ctgagaggcc 24900  
 cactgcgcgg tagagcgagc cgatctcctg gcggttcaac ggccggatcg cccccggcg 24960  
 ctcgttggtc agccgcacat cgccgatccg cgtccgcacc aggcgctgca ccgggtagcc 25020  
 gactgtcttg agcagccgcc gcacgatgcg gttgcggccc tcgtgcaaca cgatctccac 25080  
 cagcgaccgg ttctggttca tgcgaccag cttgaaccgg tccaccttca ccagaccgtc 25140  
 ctccagctcc acgccctcgc gcagctgctg cccgaggtcc ttcgggatcg acccagcac 25200  
 ctccgccagg tacgtcttgc gcaccttgta cgacgggtgc atcagccggt tcgccagctc 25260  
 gccgtcgttg gtgatcagca gcaggccctc ggtgtcctgg tccagccgcc ccacgtggaa 25320  
 cagcttgccc tcccgttccc gcaggtaatc accgatgcac ggacgcccct tgcgtccga 25380  
 catggtgcac aggatgccgc gcggttgtt cagcagcagg tgcgtcagat cgtcgttgac 25440  
 catcaccgg ttgcgtcca cgtggaccac cgcggtcgtc gggtcgacce ggcgcccag 25500  
 ctcggtgacg acctcgccgt cgacctcgat ccgcccctcg acgatcatct cctcgccgcg 25560  
 gcgcccgcgag gccaccccg cctgggacaa cactttctgc agccgcacgc cttccgaatc 25620  
 ggcagtgtg ctgtacgggg aacggtgctc agacgtcatc aatcgaatcc acttcgggca 25680

acaagggggc gagcgggcgc aactccttca acgacgacag cccagcgccg tccaggaaca 25740  
 gctcggtcgt gcagtacagg atgccgcccg tctccgggtc ggttccggcc tcttcgataa 25800  
 ggccgcgccc caccagggta cggatgacgc cgtcgcagctt cacaccccgt actgccgcga 25860  
 cccgcgaacg cgtcacccgc tgccgggtacg cgatcacccg gaggggtctcc agcgcggcac 25920  
 gcgtcagctt cgaacgctgc ccgtccaaca ggtaccgctc cacgtacggg gcgtacttct 25980  
 cccgcgtgta gaaccgccag ccgtcgcga cccgcgcgag gtcaacgccc cgttccgcct 26040  
 cggcgtagcc gtccgagagc cgccgcagcg cggaccggat ccggggccacc ggctgctcca 26100  
 ggggtgctggc cagcagctcc tcgcccggcg gcacgtcgac cacgagcagc aacgcctcca 26160  
 gcgccgcata cagcgccgaa tcggagggtca ggtccggcgc actcggcgcc tccaccgcct 26220  
 cggaatcccc agcggccgcc gacgggtcca acggttctcg cggctcagcg ggcggcgggg 26280  
 tcggcgactc ctccgcgacc tcgcccgcgg cctccggctc cagctccgtc acccgtaact 26340  
 ctcttcttca tggctggcgc gctgcgcctc ggttccggcc tccgcctgcg ccacgctgcc 26400  
 gccacccag ctacgatca gctcaccag cgggggtctcc tgctcgaaca gcaggacctt 26460  
 ctcccggtag agctcgagca gcgccaggaa ccgcgccacc acctcgacgg tgtgcccgca 26520  
 gtccgacacc agctccgaga acgtcgcgct gcccttctcc gccaacgcga cccgcagcag 26580  
 cgccgcgtgc tcgcgcaccg aaacgccgtg ctgggtggatg tgggtccagcg aaacggctcg 26640  
 cggtggcttc ggccggaaca ccgcggccgc gatctcggcg aaccgctgcg gcggaacccc 26700  
 gatcaccact tccggcagca ggttcgcgaa ccgctcctcg accgacaccg aacgcggata 26760  
 gcgccgcagc gccccggcct ccaactcgcc gaacagcgcg gccacctgct tgtacgcgcg 26820  
 gtactgcagc agccgcgcga agagcaggtc gcgggcctcc agcagcgcca ggtccgcctc 26880  
 gtctcgcacc tccgcggccg gcagcagccg ggccgccttc aggtctaaca gcgtcgcggc 26940  
 gaccaccagg aactcgggtg tttcgttcag gtcccaccgc tcgcccagcg ccttgggtgta 27000  
 ggcgatgaac tcgtcgggtga ccttgtgcag cgctacttcg gtaacgtcca gctgggtgctg 27060  
 cgagatcagc tgcagcagca ggtcgaaggg gccctcgaag ttgtccagcc gcaccgtgaa 27120  
 ccggccgcct gcggcgggat cctcgcctga ctgcgcgcgc ggtaccgcgg ggtctgtcac 27180  
 tctccgggct cctgcaggcg ccgcaccagc accgagtcgg cgcggttctt gtcgaggtcg 27240  
 gcgagcacca cggcaaccgc ttcgcgcacc acgcggcctc ggtccacggc gatgccgtgg 27300  
 tcggccccga atgccagccg cgcctgctcc aagcccagca actcctcgtc ggacacgtac 27360  
 acggtgatct tcgaatcgtg cttgcgcgct ccagagccgc cccggcccgc cgcgggtgcc 27420  
 gcctgctcgg ggtccggcgg attcgcgtga tccgtccgag gaacagtagt acggaacaat 27480  
 tcggccgctc cgggcagggc gaccgcagcg gtcacggggc gatcacctcg cgcgccagtg 27540  
 cgcggtacgc ctgcgccccg gccgagcggg gcgcccagcg ggtgatcggc tcgcccggcca 27600  
 ccgtggcttc cgaaaaccgc accgtgcggg tgatcacccg atcgaagacg atgtcgccga 27660  
 cgccctcgac gacgcgcgcc atcacctcgc gcgagtgcag cgtgcgcggg tcgaacatgg 27720  
 tcgccaggat gccgctgata tccagcctcg ggttcagccg ctcccgcacc ttctcgatgg 27780  
 tgctgatcag cagcgccacc ccgcgcaggc tgaagaactc gcaactccagc gggattatca 27840  
 cggcgctccg ggccggccaa gcgttcaccg tcagcagccc cagcgacggc tggcagtcca 27900  
 ccagcacgta gtcgtactcc gccagcgccg gctgcagcac ccggtgcagc gtctgctccc 27960  
 ggcccacctc cgcgaccagc tgcacctcgg ccgcggacag gtcgatgttg ctgggcagca 28020  
 ggtccatgcc ctgcagcgtg gtctgcgcga ccacgtcatg gatggcgacc ggccgttcca 28080  
 tgatcacgtt gtagatggtc tggctcagct ggtgcggctg gaccccagag ccgaccgaga 28140  
 gcgcgccttg cggatcgaaa tccaccagca gcacccggcg cccgtattcc gccagcgctg 28200  
 cgccgaggtt gatcgtcgac gtggtcttcc cgacgcgcgc cttctggttg cacatcgcca 28260  
 gcaccgatgc gggaccgtgc cgggtccagca gcggcggtc ggggatgtgc cgacgtggcc 28320  
 ggccagtggg accgaggccg cgggggttcg gcgccaggcc gacctcctcc tcgcccggcg 28380  
 ccgtctcccg ggcgatgctc aggtcgaccg cgccccggga ccgcccttcg gcggcgggct 28440  
 gcggcagcga catggcgatc tgactccttt tgggtgccgt ggcgatcaaa cgcagcctaa 28500  
 gtggcattcg ttaccagcgg caacgcgcct cgccgtgggt tcggaagtcc cctgtgatct 28560

```

gcgtgtgata aacgtctacc gcgtacacca cccctcgcga ttgcgagcgt caccaagtca 28620
ctccgtagcg tgtgcgcgcg gatgcgtcgt cgcgtagacc tcccgcagcg tggtcaccgt 28680
gatcaaggtg tagacctggg tcgtcgtgac cgaggcgtgg cccagcagtt cctgcaccac 28740
gcggacatcc gcgccgccct ccacaaggtg cgtggcgaac gaatgcgcga gcacgtgcgg 28800
cgaaaccgag ccgttgatcc cggcgcgctg cgcggcggtc ttcaacgcgt tccacgagct 28860
ctgccgggaa agccgggtac cgcgcgaatt caggaacacc gcggcactac cccgcccgcg 28920
cgcggccagg ccgggcctgg cccgtaccag gtacgcgtcc agcgcgcgca gcgccggtcg 28980
gccgatgggc accagccgct ggccggccgc cttcccttcc agcagcacgg tccggttggt 29040
gccgtcgatg tcgtcgaggt ccaaccgcac cgcttcggag atccgcgctc ccgtcgaata 29100
cagcacttcc agcagcgccc ggtcccgcag gccccgggca tctcgcctcc cggcgtgatc 29160
gagcagcttc ccgacgtcgc cgatggaaag cgcttcggc agccgcttcg gcgggctcgg 29220
cggcgcgaca tcgcgcgcta cgtcgaccgg cagcagcccc tcggcgtgcg cgaagcgggtg 29280
caggccgcgg gcggccacca aagctcgggc cgacgacgac ggccgcagcg cgggcccgtg 29340
cccgggtgct tcccggagct ccgccaagaa cccggacagg tgctccgaac gcacctcgcc 29400
gagcccagaa atccctgccg cgaccaggaa ttcggcatac cggcgcaggt ccctggcgta 29460
ggagtcgaga gtgctgcggg cggtagcgcg ctcgacggca aggtggtcga ggtaccgggt 29520
gatcgcctcg cgcaggtccg gcggcagctc ctcgaaacct ctcaccgggc gggaagcgc 29580
ggggccgggt ccggccattc ggcacccgcc gggcgaggct ccaccggcc ctcgcggaac 29640
gcgtgggcgg cgagcagccc cgccaccccg gggccggtga cgatctcccc ggccagcgcc 29700
atccgcaccg cctcggcgag cgggaagcgg cggatc 29736

```

<210> 7

<211> 828

<212> DNA

<213> *Saccharopolyspora spinosa*

<220>

<221> CDS

<222> (1)..(825)

<223> ORF1; O-Methyltransferase

<400> 7

```

tg ttg cca ggt ggc gca cca aca tcg cag cag gtt ggg cag atg tat 48
Val Leu Pro Gly Gly Ala Pro Thr Ser Gln Gln Val Gly Gln Met Tyr
1 5 10 15

gac ctg gtc acg ccg ttg ctg aac tcg gtc gcg ggc ggc ccc tgc gcc 96
Asp Leu Val Thr Pro Leu Leu Asn Ser Val Ala Gly Gly Pro Cys Ala
20 25 30

atc cac cac ggc tac tgg gag aac gac ggg cgg gct tcc tgg cag cag 144
Ile His His Gly Tyr Trp Glu Asn Asp Gly Arg Ala Ser Trp Gln Gln
35 40 45

gcc gcc gac cgg ctc acc gac ctt gtc gcc gaa cgg acc gtg ctc gat 192
Ala Ala Asp Arg Leu Thr Asp Leu Val Ala Glu Arg Thr Val Leu Asp
50 55 60

```

ggc ggc gtt cga ctg ctc gat gtg ggg tgc ggt acc gga caa cca gcg	240
Gly Gly Val Arg Leu Leu Asp Val Gly Cys Gly Thr Gly Gln Pro Ala	
65 70 75 80	
ctg cgc gtc gcg cgc gac aac gcg atc cag atc acc ggc atc acc gtc	288
Leu Arg Val Ala Arg Asp Asn Ala Ile Gln Ile Thr Gly Ile Thr Val	
85 90 95	
agc cag gtg caa gtg gcc atc gcc gct gat tgc gca cgc gaa cgc gga	336
Ser Gln Val Gln Val Ala Ile Ala Ala Asp Cys Ala Arg Glu Arg Gly	
100 105 110	
cta agc cac cgg gtg gac ttc tcg tgc gtc gat gcc atg tcc ctg ccg	384
Leu Ser His Arg Val Asp Phe Ser Cys Val Asp Ala Met Ser Leu Pro	
115 120 125	
tac ccg gac aat gct ttc gac gcc gcc tgg gcc atg cag tcg ctg ttg	432
Tyr Pro Asp Asn Ala Phe Asp Ala Ala Trp Ala Met Gln Ser Leu Leu	
130 135 140	
gag atg tcc gaa ccg gac cgt gcc atc cgg gaa atc ctt cga gta ctc	480
Glu Met Ser Glu Pro Asp Arg Ala Ile Arg Glu Ile Leu Arg Val Leu	
145 150 155 160	
aaa ccc ggt ggc atc ctc ggc gtc acc gag gtc gtc aaa cga gaa gcg	528
Lys Pro Gly Gly Ile Leu Gly Val Thr Glu Val Val Lys Arg Glu Ala	
165 170 175	
ggc ggc ggg atg ccg gtg tcc ggg gac agg tgg ccg acc ggc ctt ccg	576
Gly Gly Gly Met Pro Val Ser Gly Asp Arg Trp Pro Thr Gly Leu Arg	
180 185 190	
atc tgc ctg gct gag caa ctt ctg gaa tcg ctg cgt gca gcg ggg ttc	624
Ile Cys Leu Ala Glu Gln Leu Leu Glu Ser Leu Arg Ala Ala Gly Phe	
195 200 205	
gag atc ctc gat tgg gag gac gtg tcg tcg agg acc ccg tac ttc atg	672
Glu Ile Leu Asp Trp Glu Asp Val Ser Ser Arg Thr Arg Tyr Phe Met	
210 215 220	
ccg cag ttc gcc gaa gag ctc gct gcg cac cag cac ggg atc gcg gac	720
Pro Gln Phe Ala Glu Glu Leu Ala Ala His Gln His Gly Ile Ala Asp	
225 230 235 240	
agg tac ggg ccg gct gtc gcc ggc tgg gcc gcc gcg gtc tgc gat tat	768
Arg Tyr Gly Pro Ala Val Ala Gly Trp Ala Ala Ala Val Cys Asp Tyr	
245 250 255	

gag aaa tat gcc cac gac atg ggc tat gcg att ctg acg gcg cgg aag 816  
 Glu Lys Tyr Ala His Asp Met Gly Tyr Ala Ile Leu Thr Ala Arg Lys  
                   260                                  265                                  270

ccg gtc ggc tga 828  
 Pro Val Gly  
                   275

<210> 8  
 <211> 275  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 8  
 Val Leu Pro Gly Gly Ala Pro Thr Ser Gln Gln Val Gly Gln Met Tyr  
   1                                  5                                  10                                  15

Asp Leu Val Thr Pro Leu Leu Asn Ser Val Ala Gly Gly Pro Cys Ala  
                   20                                  25                                  30

Ile His His Gly Tyr Trp Glu Asn Asp Gly Arg Ala Ser Trp Gln Gln  
                   35                                  40                                  45

Ala Ala Asp Arg Leu Thr Asp Leu Val Ala Glu Arg Thr Val Leu Asp  
                   50                                  55                                  60

Gly Gly Val Arg Leu Leu Asp Val Gly Cys Gly Thr Gly Gln Pro Ala  
   65                                  70                                  75                                  80

Leu Arg Val Ala Arg Asp Asn Ala Ile Gln Ile Thr Gly Ile Thr Val  
                                   85                                  90                                  95

Ser Gln Val Gln Val Ala Ile Ala Ala Asp Cys Ala Arg Glu Arg Gly  
                   100                                  105                                  110

Leu Ser His Arg Val Asp Phe Ser Cys Val Asp Ala Met Ser Leu Pro  
                   115                                  120                                  125

Tyr Pro Asp Asn Ala Phe Asp Ala Ala Trp Ala Met Gln Ser Leu Leu  
                   130                                  135                                  140

Glu Met Ser Glu Pro Asp Arg Ala Ile Arg Glu Ile Leu Arg Val Leu  
   145                                  150                                  155                                  160

Lys Pro Gly Gly Ile Leu Gly Val Thr Glu Val Val Lys Arg Glu Ala  
                   165                                  170                                  175

Gly Gly Gly Met Pro Val Ser Gly Asp Arg Trp Pro Thr Gly Leu Arg  
 180 185 190

Ile Cys Leu Ala Glu Gln Leu Leu Glu Ser Leu Arg Ala Ala Gly Phe  
 195 200 205

Glu Ile Leu Asp Trp Glu Asp Val Ser Ser Arg Thr Arg Tyr Phe Met  
 210 215 220

Pro Gln Phe Ala Glu Glu Leu Ala Ala His Gln His Gly Ile Ala Asp  
 225 230 235 240

Arg Tyr Gly Pro Ala Val Ala Gly Trp Ala Ala Ala Val Cys Asp Tyr  
 245 250 255

Glu Lys Tyr Ala His Asp Met Gly Tyr Ala Ile Leu Thr Ala Arg Lys  
 260 265 270

Pro Val Gly  
 275

<210> 9

<211> 1173

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1170)

<223> ORF2; Glycosyltransferase

<400> 9

atg cgc gta ctc gtc gtt ccc ttg ccc tat ccg acg cat ctc atg gca 48  
 Met Arg Val Leu Val Val Pro Leu Pro Tyr Pro Thr His Leu Met Ala  
 1 5 10 15

atg gtg ccg ctg tgc tgg gcg ctg caa gca tcc ggg cac gag gtc ctg 96  
 Met Val Pro Leu Cys Trp Ala Leu Gln Ala Ser Gly His Glu Val Leu  
 20 25 30

atc gcc gca cca cca gag ctg cag gcg acc gcg cat ggt gca ggt ctc 144  
 Ile Ala Ala Pro Pro Glu Leu Gln Ala Thr Ala His Gly Ala Gly Leu  
 35 40 45

acc acg gcc ggg atc cgc ggg aac gac agg acc ggc gat acg ggt gga 192

Thr Thr Ala Gly Ile Arg Gly Asn Asp Arg Thr Gly Asp Thr Gly Gly	
50 55 60	
acc acg cag ctg cgc ttt ccc aat ccg gcg ttc ggt cag cgc gac acc	240
Thr Thr Gln Leu Arg Phe Pro Asn Pro Ala Phe Gly Gln Arg Asp Thr	
65 70 75 80	
gag gca ggc cgg caa ctg tgg gag cag acc gcg tcc aat gtc gcg caa	288
Glu Ala Gly Arg Gln Leu Trp Glu Gln Thr Ala Ser Asn Val Ala Gln	
85 90 95	
agc tcg ctc gat cag ctc ccc gaa tac ctt cga ctg gcc gag gcc tgg	336
Ser Ser Leu Asp Gln Leu Pro Glu Tyr Leu Arg Leu Ala Glu Ala Trp	
100 105 110	
cga ccg tca gtg ctg ttg gtc gac gtc tgc gcg ctg atc ggc cgg gtg	384
Arg Pro Ser Val Leu Leu Val Asp Val Cys Ala Leu Ile Gly Arg Val	
115 120 125	
ctc ggc gga ttg ctc gac ctg ccg gtc gtg ctg cac cgc tgg gga gtc	432
Leu Gly Gly Leu Leu Asp Leu Pro Val Val Leu His Arg Trp Gly Val	
130 135 140	
gac ccc acc gca ggc ccc ttc agc gat cga gcc cac gag ttg ctt gac	480
Asp Pro Thr Ala Gly Pro Phe Ser Asp Arg Ala His Glu Leu Leu Asp	
145 150 155 160	
ccg gtg tgc cgg cac cac gga ctg acc ggc ctg ccc act ccc gag ctc	528
Pro Val Cys Arg His His Gly Leu Thr Gly Leu Pro Thr Pro Glu Leu	
165 170 175	
atc ctc gat ccc tgt ccg ccg agc ctg caa gca agc gac gcg ccg caa	576
Ile Leu Asp Pro Cys Pro Pro Ser Leu Gln Ala Ser Asp Ala Pro Gln	
180 185 190	
ggc gca ccg gtc cag tac gtg ccg tac aac gga agc ggc gca ttc ccg	624
Gly Ala Pro Val Gln Tyr Val Pro Tyr Asn Gly Ser Gly Ala Phe Pro	
195 200 205	
gca tgg ggc gcg gcg cgc acc tca gca cgg cgg gtc tgc atc tgc atg	672
Ala Trp Gly Ala Ala Arg Thr Ser Ala Arg Arg Val Cys Ile Cys Met	
210 215 220	
ggc cgc atg gtg ctg aac gcc acc ggg ccg gct ccg ctg ctg cgc gca	720
Gly Arg Met Val Leu Asn Ala Thr Gly Pro Ala Pro Leu Leu Arg Ala	
225 230 235 240	
gta gcg gct gcc acc gag ttg ccc ggc gtc gag gcc gtg atc gcc gtt	768



Val Ala Ala Ala Thr Glu Leu Pro Gly Val Glu Ala Val Ile Ala Val  
 245 250 255

ccc ccc gag cac cgg gca ctt ctc acc gac cta ccc gac aac gcc cgg 816  
 Pro Pro Glu His Arg Ala Leu Leu Thr Asp Leu Pro Asp Asn Ala Arg  
 260 265 270

atc gcc gaa tcg gtc ccg ctc aac ctg ttc ctg cgt acc tgc gag ctg 864  
 Ile Ala Glu Ser Val Pro Leu Asn Leu Phe Leu Arg Thr Cys Glu Leu  
 275 280 285

gtc atc tgc gcg ggc ggc tcg gga acg gca ttc acc gcg acc cga ctc 912  
 Val Ile Cys Ala Gly Gly Ser Gly Thr Ala Phe Thr Ala Thr Arg Leu  
 290 295 300

ggc atc ccg caa ctc gtg ctt ccc cag tac ttc gac cag ttc gac tac 960  
 Gly Ile Pro Gln Leu Val Leu Pro Gln Tyr Phe Asp Gln Phe Asp Tyr  
 305 310 315 320

gcg cgc aac ctc gcc gct gcc ggg gcg ggc atc tgc ttg ccg gat gag 1008  
 Ala Arg Asn Leu Ala Ala Ala Gly Ala Gly Ile Cys Leu Pro Asp Glu  
 325 330 335

cag gcc cag tcc gac cac gaa cag ttc acc gac tca atc gca acg gtg 1056  
 Gln Ala Gln Ser Asp His Glu Gln Phe Thr Asp Ser Ile Ala Thr Val  
 340 345 350

ctc ggc gac acc ggc ttc gct gct gcg gca atc aaa ctc agc gac gag 1104  
 Leu Gly Asp Thr Gly Phe Ala Ala Ala Ile Lys Leu Ser Asp Glu  
 355 360 365

atc acg gcc atg ccc cat ccc gcc gcg ctg gtg cgg acg ctg gag aac 1152  
 Ile Thr Ala Met Pro His Pro Ala Ala Leu Val Arg Thr Leu Glu Asn  
 370 375 380

act gcg gcc atc cgt gcc tga 1173  
 Thr Ala Ala Ile Arg Ala  
 385 390

<210> 10

<211> 390

<212> PRT

<213> Saccharopolyspora spinosa

<400> 10

Met Arg Val Leu Val Val Pro Leu Pro Tyr Pro Thr His Leu Met Ala  
 1 5 10 15

Met Val Pro Leu Cys Trp Ala Leu Gln Ala Ser Gly His Glu Val Leu  
20 25 30

Ile Ala Ala Pro Pro Glu Leu Gln Ala Thr Ala His Gly Ala Gly Leu  
35 40 45

Thr Thr Ala Gly Ile Arg Gly Asn Asp Arg Thr Gly Asp Thr Gly Gly  
50 55 60

Thr Thr Gln Leu Arg Phe Pro Asn Pro Ala Phe Gly Gln Arg Asp Thr  
65 70 75 80

Glu Ala Gly Arg Gln Leu Trp Glu Gln Thr Ala Ser Asn Val Ala Gln  
85 90 95

Ser Ser Leu Asp Gln Leu Pro Glu Tyr Leu Arg Leu Ala Glu Ala Trp  
100 105 110

Arg Pro Ser Val Leu Leu Val Asp Val Cys Ala Leu Ile Gly Arg Val  
115 120 125

Leu Gly Gly Leu Leu Asp Leu Pro Val Val Leu His Arg Trp Gly Val  
130 135 140

Asp Pro Thr Ala Gly Pro Phe Ser Asp Arg Ala His Glu Leu Leu Asp  
145 150 155 160

Pro Val Cys Arg His His Gly Leu Thr Gly Leu Pro Thr Pro Glu Leu  
165 170 175

Ile Leu Asp Pro Cys Pro Pro Ser Leu Gln Ala Ser Asp Ala Pro Gln  
180 185 190

Gly Ala Pro Val Gln Tyr Val Pro Tyr Asn Gly Ser Gly Ala Phe Pro  
195 200 205

Ala Trp Gly Ala Ala Arg Thr Ser Ala Arg Arg Val Cys Ile Cys Met  
210 215 220

Gly Arg Met Val Leu Asn Ala Thr Gly Pro Ala Pro Leu Leu Arg Ala  
225 230 235 240

Val Ala Ala Ala Thr Glu Leu Pro Gly Val Glu Ala Val Ile Ala Val  
245 250 255

Pro Pro Glu His Arg Ala Leu Leu Thr Asp Leu Pro Asp Asn Ala Arg  
260 265 270

Ile Ala Glu Ser Val Pro Leu Asn Leu Phe Leu Arg Thr Cys Glu Leu  
275 280 285

Val Ile Cys Ala Gly Gly Ser Gly Thr Ala Phe Thr Ala Thr Arg Leu  
290 295 300

Gly Ile Pro Gln Leu Val Leu Pro Gln Tyr Phe Asp Gln Phe Asp Tyr  
305 310 315 320

Ala Arg Asn Leu Ala Ala Ala Gly Ala Gly Ile Cys Leu Pro Asp Glu  
325 330 335

Gln Ala Gln Ser Asp His Glu Gln Phe Thr Asp Ser Ile Ala Thr Val  
340 345 350

Leu Gly Asp Thr Gly Phe Ala Ala Ala Ala Ile Lys Leu Ser Asp Glu  
355 360 365

Ile Thr Ala Met Pro His Pro Ala Ala Leu Val Arg Thr Leu Glu Asn  
370 375 380

Thr Ala Ala Ile Arg Ala  
385 390

<210> 11

<211> 753

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(750)

<223> ORF3; O-Methyltransferase

<400> 11

atg ccc tcc cag aac gcg ctg tac ctg gac ctg ctc aag aag gta ctc 48  
Met Pro Ser Gln Asn Ala Leu Tyr Leu Asp Leu Leu Lys Lys Val Leu  
1 5 10 15

acc aac acg att tac agt gat cgg ccg cat ccg aac gcc tgg cag gac 96  
Thr Asn Thr Ile Tyr Ser Asp Arg Pro His Pro Asn Ala Trp Gln Asp  
20 25 30

aac acc gac tac agg cag gcc gct cgg gcc aaa ggc acg gac tgg cca 144  
Asn Thr Asp Tyr Arg Gln Ala Ala Arg Ala Lys Gly Thr Asp Trp Pro

act gtc gcg cac acg atg atc ggt ctg gag cgg ctg gac aac ctc cag 192  
 Thr Val Ala His Thr Met Ile Gly Leu Glu Arg Leu Asp Asn Leu Gln  
 50 55 60

cac tgc gtg gaa gcc gtg ctc gca gac ggt gtt ccc ggg gat ttc gcc 240  
 His Cys Val Glu Ala Val Leu Ala Asp Gly Val Pro Gly Asp Phe Ala  
 65 70 75 80

gag acc ggt gtc tgg cgg ggc ggc gca tgc atc ttc atg cgc gcg gtt 288  
 Glu Thr Gly Val Trp Arg Gly Gly Ala Cys Ile Phe Met Arg Ala Val  
 85 90 95

ctc cag gca ttc gga gat acc gga cgt acc gtc tgg gta gtg gat tcc 336  
 Leu Gln Ala Phe Gly Asp Thr Gly Arg Thr Val Trp Val Val Asp Ser  
 100 105 110

ttc cag gga atg ccg gaa agc tct gcg caa gac cac caa gcg gac cag 384  
 Phe Gln Gly Met Pro Glu Ser Ser Ala Gln Asp His Gln Ala Asp Gln  
 115 120 125

gct atg gcg ctg cac gag tac aac gac gtg ctt ggc gta tcg ctt gag 432  
 Ala Met Ala Leu His Glu Tyr Asn Asp Val Leu Gly Val Ser Leu Glu  
 130 135 140

acc gtc cgg cag aac ttc gcc cgc tac ggg ctg ctc gac gaa cag gtc 480  
 Thr Val Arg Gln Asn Phe Ala Arg Tyr Gly Leu Leu Asp Glu Gln Val  
 145 150 155 160

agg ttc ctc ccc ggc tgg ttc cgg gac acc ttg ccc acc gcc ccc atc 528  
 Arg Phe Leu Pro Gly Trp Phe Arg Asp Thr Leu Pro Thr Ala Pro Ile  
 165 170 175

cag gaa ctc gcc gtg cta cga ctc gac ggc gac ctc tac gaa tcc aca 576  
 Gln Glu Leu Ala Val Leu Arg Leu Asp Gly Asp Leu Tyr Glu Ser Thr  
 180 185 190

atg gac tca ttg cgg aac ctg tac ccg aag ctc tcg ccg ggc gga ttc 624  
 Met Asp Ser Leu Arg Asn Leu Tyr Pro Lys Leu Ser Pro Gly Gly Phe  
 195 200 205

gtc atc atc gac gac tat ttt ctg ccg tcc tgc cag gac gcg gtg aag 672  
 Val Ile Ile Asp Asp Tyr Phe Leu Pro Ser Cys Gln Asp Ala Val Lys  
 210 215 220

ggg ttc cgc gcg gaa ctc ggg atc acg gaa ccc atc cac gac atc gac 720  
 Gly Phe Arg Ala Glu Leu Gly Ile Thr Glu Pro Ile His Asp Ile Asp

225

230

235

240

ggc acg ggc gcc tac tgg cgc cgc agc tgg tga  
 Gly Thr Gly Ala Tyr Trp Arg Arg Ser Trp  
 245 250

753

&lt;210&gt; 12

&lt;211&gt; 250

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 12

Met Pro Ser Gln Asn Ala Leu Tyr Leu Asp Leu Leu Lys Lys Val Leu  
 1 5 10 15

Thr Asn Thr Ile Tyr Ser Asp Arg Pro His Pro Asn Ala Trp Gln Asp  
 20 25 30

Asn Thr Asp Tyr Arg Gln Ala Ala Arg Ala Lys Gly Thr Asp Trp Pro  
 35 40 45

Thr Val Ala His Thr Met Ile Gly Leu Glu Arg Leu Asp Asn Leu Gln  
 50 55 60

His Cys Val Glu Ala Val Leu Ala Asp Gly Val Pro Gly Asp Phe Ala  
 65 70 75 80

Glu Thr Gly Val Trp Arg Gly Gly Ala Cys Ile Phe Met Arg Ala Val  
 85 90 95

Leu Gln Ala Phe Gly Asp Thr Gly Arg Thr Val Trp Val Val Asp Ser  
 100 105 110

Phe Gln Gly Met Pro Glu Ser Ser Ala Gln Asp His Gln Ala Asp Gln  
 115 120 125

Ala Met Ala Leu His Glu Tyr Asn Asp Val Leu Gly Val Ser Leu Glu  
 130 135 140

Thr Val Arg Gln Asn Phe Ala Arg Tyr Gly Leu Leu Asp Glu Gln Val  
 145 150 155 160

Arg Phe Leu Pro Gly Trp Phe Arg Asp Thr Leu Pro Thr Ala Pro Ile  
 165 170 175

Gln Glu Leu Ala Val Leu Arg Leu Asp Gly Asp Leu Tyr Glu Ser Thr  
 180 185 190

Met Asp Ser Leu Arg Asn Leu Tyr Pro Lys Leu Ser Pro Gly Gly Phe  
 195 200 205

Val Ile Ile Asp Asp Tyr Phe Leu Pro Ser Cys Gln Asp Ala Val Lys  
 210 215 220

Gly Phe Arg Ala Glu Leu Gly Ile Thr Glu Pro Ile His Asp Ile Asp  
 225 230 235 240

Gly Thr Gly Ala Tyr Trp Arg Arg Ser Trp  
 245 250

<210> 13

<211> 1188

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1185)

<223> ORF4; O-Methyltransferase

<400> 13

atg agt gag atc gca gtt gcc ccc tgg tcg gtg gtg gag cgt ttg ctg 48  
 Met Ser Glu Ile Ala Val Ala Pro Trp Ser Val Val Glu Arg Leu Leu  
 1 5 10 15

ctc gcg gcg ggt gcg ggc ccg gcg aag ctc cag gaa gca gtg cag gtg 96  
 Leu Ala Ala Gly Ala Gly Pro Ala Lys Leu Gln Glu Ala Val Gln Val  
 20 25 30

gcc gga ctg gac gcg gtg gcc gac gcc atc gtc gac gaa ctc gtc gta 144  
 Ala Gly Leu Asp Ala Val Ala Asp Ala Ile Val Asp Glu Leu Val Val  
 35 40 45

cgc tgc gat ccg ctg tcg ttg gac gag tcg gtg cga atc ggc ctg gag 192  
 Arg Cys Asp Pro Leu Ser Leu Asp Glu Ser Val Arg Ile Gly Leu Glu  
 50 55 60

atc act tct ggc gct cag ctg gtc cgg aga acc gtt gag ctc gat cac 240  
 Ile Thr Ser Gly Ala Gln Leu Val Arg Arg Thr Val Glu Leu Asp His  
 65 70 75 80

gca ggc ctg cgg ctc gcg gcg gtc gcc gaa gca gct gct gtt ctc cgg 288  
 Ala Gly Leu Arg Leu Ala Ala Val Ala Glu Ala Ala Val Leu Arg

ttc gac gcg gtg gat ctg ctg gaa ggg ctc ttc ggc ccg gtt gac ggc	336
Phe Asp Ala Val Asp Leu Leu Glu Gly Leu Phe Gly Pro Val Asp Gly	
100 105 110	
agg cgg cac aac agc cgt gaa gtc cgc tgg tcg gac agc atg acg cag	384
Arg Arg His Asn Ser Arg Glu Val Arg Trp Ser Asp Ser Met Thr Gln	
115 120 125	
ttc tcg ccc gac cag ggc ctc gcc ggc gcg cag cgc ctg ctg gcg ttc	432
Phe Ser Pro Asp Gln Gly Leu Ala Gly Ala Gln Arg Leu Leu Ala Phe	
130 135 140	
cgg aac agg gtg tcc acc gcg gtg cac gcc gtg ctg gcc gca gcc gcc	480
Arg Asn Arg Val Ser Thr Ala Val His Ala Val Leu Ala Ala Ala Ala	
145 150 155 160	
acc agg cgc gcg gac ctc ggt gcg ctg gca gtc cgc tac gga tcc gac	528
Thr Arg Arg Ala Asp Leu Gly Ala Leu Ala Val Arg Tyr Gly Ser Asp	
165 170 175	
aaa tgg gcg gac ctg cac tgg tac acc gaa cac tac gag cac cac ttc	576
Lys Trp Ala Asp Leu His Trp Tyr Thr Glu His Tyr Glu His His Phe	
180 185 190	
tcc cga ttc cag gat gcc ccg gtg cga gtg ttg gaa ata gga atc ggt	624
Ser Arg Phe Gln Asp Ala Pro Val Arg Val Leu Glu Ile Gly Ile Gly	
195 200 205	
ggt tat cac gca ccc gaa ctc ggt ggt gct tcg ctg cgc atg tgg cag	672
Gly Tyr His Ala Pro Glu Leu Gly Gly Ala Ser Leu Arg Met Trp Gln	
210 215 220	
cgg tac ttc cgg cga ggt ctc gtt tac ggg ctg gac att ttc gag aaa	720
Arg Tyr Phe Arg Arg Gly Leu Val Tyr Gly Leu Asp Ile Phe Glu Lys	
225 230 235 240	
gcc ggg aac gaa ggg cac cga gtg cga aag ctg cga ggt gac cag agc	768
Ala Gly Asn Glu Gly His Arg Val Arg Lys Leu Arg Gly Asp Gln Ser	
245 250 255	
gat gcg gaa ttc ctg gaa gac atg gcg ggg aag atc ggg ccg ttc gac	816
Asp Ala Glu Phe Leu Glu Asp Met Ala Gly Lys Ile Gly Pro Phe Asp	
260 265 270	
att gtc atc gac gac ggc agc cat gtc aac gac cac gtc aag aaa tcc	864
Ile Val Ile Asp Asp Gly Ser His Val Asn Asp His Val Lys Lys Ser	

275

280

2

ttc caa tcc ctg ttt ccg cac gtc cgc cca ggt ggt ttg tac gtc atc 912  
 Phe Gln Ser Leu Phe Pro His Val Arg Pro Gly Gly Leu Tyr Val Ile  
 290 295 300

gag gat ctc cag acg gcg tac tgg ccc ggc tac ggc ggt cgc gat ggg 960  
 Glu Asp Leu Gln Thr Ala Tyr Trp Pro Gly Tyr Gly Gly Arg Asp Gly  
 305 310 315 320

gaa ccc gcg gcc cag cgc acc tcg atc gac atg ctc aaa gaa ctg atc 1008  
 Glu Pro Ala Ala Gln Arg Thr Ser Ile Asp Met Leu Lys Glu Leu Ile  
 325 330 335

gac ggc ctg cat tat cag gag cgc gaa tcg cgg tgc ggg acc gag ccc 1056  
 Asp Gly Leu His Tyr Gln Glu Arg Glu Ser Arg Cys Gly Thr Glu Pro  
 340 345 350

tcc tac acg gaa cgg aac gtg gcg gcc ctg cac ttc tac cac aac ctg 1104  
 Ser Tyr Thr Glu Arg Asn Val Ala Ala Leu His Phe Tyr His Asn Leu  
 355 360 365

gta ttc gtg gag aaa ggg ctc aac gct gag cct gcc gcg ccg ggg ttc 1152  
 Val Phe Val Glu Lys Gly Leu Asn Ala Glu Pro Ala Ala Pro Gly Phe  
 370 375 380

gtg ccc cgg caa gcg ctc ggc gtc gag gac ggc tga 1188  
 Val Pro Arg Gln Ala Leu Gly Val Glu Asp Gly  
 385 390 395

&lt;210&gt; 14

&lt;211&gt; 395

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 14

Met Ser Glu Ile Ala Val Ala Pro Trp Ser Val Val Glu Arg Leu Leu  
 1 5 10 15

Leu Ala Ala Gly Ala Gly Pro Ala Lys Leu Gln Glu Ala Val Gln Val  
 20 25 30

Ala Gly Leu Asp Ala Val Ala Asp Ala Ile Val Asp Glu Leu Val Val  
 35 40 45

Arg Cys Asp Pro Leu Ser Leu Asp Glu Ser Val Arg Ile Gly Leu Glu  
 50 55 60



Ile Thr Ser Gly Ala Gln Leu Val Arg Arg Thr Val Glu Leu Asp His  
 65 70 75 80  
 Ala Gly Leu Arg Leu Ala Ala Val Ala Glu Ala Ala Ala Val Leu Arg  
 85 90 95  
 Phe Asp Ala Val Asp Leu Leu Glu Gly Leu Phe Gly Pro Val Asp Gly  
 100 105 110  
 Arg Arg His Asn Ser Arg Glu Val Arg Trp Ser Asp Ser Met Thr Gln  
 115 120 125  
 Phe Ser Pro Asp Gln Gly Leu Ala Gly Ala Gln Arg Leu Leu Ala Phe  
 130 135 140  
 Arg Asn Arg Val Ser Thr Ala Val His Ala Val Leu Ala Ala Ala Ala  
 145 150 155 160  
 Thr Arg Arg Ala Asp Leu Gly Ala Leu Ala Val Arg Tyr Gly Ser Asp  
 165 170 175  
 Lys Trp Ala Asp Leu His Trp Tyr Thr Glu His Tyr Glu His His Phe  
 180 185 190  
 Ser Arg Phe Gln Asp Ala Pro Val Arg Val Leu Glu Ile Gly Ile Gly  
 195 200 205  
 Gly Tyr His Ala Pro Glu Leu Gly Gly Ala Ser Leu Arg Met Trp Gln  
 210 215 220  
 Arg Tyr Phe Arg Arg Gly Leu Val Tyr Gly Leu Asp Ile Phe Glu Lys  
 225 230 235 240  
 Ala Gly Asn Glu Gly His Arg Val Arg Lys Leu Arg Gly Asp Gln Ser  
 245 250 255  
 Asp Ala Glu Phe Leu Glu Asp Met Ala Gly Lys Ile Gly Pro Phe Asp  
 260 265 270  
 Ile Val Ile Asp Asp Gly Ser His Val Asn Asp His Val Lys Lys Ser  
 275 280 285  
 Phe Gln Ser Leu Phe Pro His Val Arg Pro Gly Gly Leu Tyr Val Ile  
 290 295 300  
 Glu Asp Leu Gln Thr Ala Tyr Trp Pro Gly Tyr Gly Gly Arg Asp Gly  
 305 310 315 320

Glu Pro Ala Ala Gln Arg Thr Ser Ile Asp Met Leu Lys Glu Leu Ile  
325 330 335

Asp Gly Leu His Tyr Gln Glu Arg Glu Ser Arg Cys Gly Thr Glu Pro  
340 345 350

Ser Tyr Thr Glu Arg Asn Val Ala Ala Leu His Phe Tyr His Asn Leu  
355 360 365

Val Phe Val Glu Lys Gly Leu Asn Ala Glu Pro Ala Ala Pro Gly Phe  
370 375 380

Val Pro Arg Gln Ala Leu Gly Val Glu Asp Gly  
385 390 395

<210> 15

<211> 1620

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1617)

<223> ORF5; C-C verknüpfendes Enzym, Cyclisierungsenzym

<400> 15

atg atc tcg gct gcg ggc gaa caa agt gga cca gtc aga aaa gga ggg 48  
Met Ile Ser Ala Ala Gly Glu Gln Ser Gly Pro Val Arg Lys Gly Gly  
1 5 10 15

gcg gtg ccc gaa ttc cat gac ccg gca ccc atg aat cgt cga acc cca 96  
Ala Val Pro Glu Phe His Asp Pro Ala Pro Met Asn Arg Arg Thr Pro  
20 25 30

gga aca gag atc acc gtc gag ccc gac gat cct cgt tat ccg gac ctc 144  
Gly Thr Glu Ile Thr Val Glu Pro Asp Asp Pro Arg Tyr Pro Asp Leu  
35 40 45

gtc gtc ggg cac aac ccc cgt ttc acc gga aaa ccc gaa cgc atc cac 192  
Val Val Gly His Asn Pro Arg Phe Thr Gly Lys Pro Glu Arg Ile His  
50 55 60

atc gcc agc tcc gcc gaa gac gtc gtg cac gcc gtc gcc gac gcc gtg 240  
Ile Ala Ser Ser Ala Glu Asp Val Val His Ala Val Ala Asp Ala Val  
65 70 75 80

cgc acc ggc agg cgg gta ggg gtc cgc agc ggc ggg cac tgc ttc gag	288
Arg Thr Gly Arg Arg Val Gly Val Arg Ser Gly Gly His Cys Phe Glu	
85 90 95	
aat ctc gtt gcg gac ccg gcg atc cga gtg ctc gtc gac ctc tcc gag	336
Asn Leu Val Ala Asp Pro Ala Ile Arg Val Leu Val Asp Leu Ser Glu	
100 105 110	
ctc aac cgc gtg tac tac gac agc acg cgc ggg gca ttc gcg atc gag	384
Leu Asn Arg Val Tyr Tyr Asp Ser Thr Arg Gly Ala Phe Ala Ile Glu	
115 120 125	
gcg ggc gcc gcc ctc ggg cag gtg tac cga acc ctg ttc aag aac tgg	432
Ala Gly Ala Ala Leu Gly Gln Val Tyr Arg Thr Leu Phe Lys Asn Trp	
130 135 140	
ggc gtg acg atc ccg acc ggc gca tgt ccc ggg gtg ggc gca ggc ggg	480
Gly Val Thr Ile Pro Thr Gly Ala Cys Pro Gly Val Gly Ala Gly Gly	
145 150 155 160	
cac atc ctc ggc ggg gga tac ggc ccg ctg tcg cgc cga ttc ggt tcg	528
His Ile Leu Gly Gly Gly Tyr Gly Pro Leu Ser Arg Arg Phe Gly Ser	
165 170 175	
gtc gtc gac tac ctt caa ggc gtc gag gtc gtc gtg gtc gac cag gcc	576
Val Val Asp Tyr Leu Gln Gly Val Glu Val Val Val Val Asp Gln Ala	
180 185 190	
ggt gaa gtg cac atc gtc gag gcc gac cgg aac tcc acg ggc gcc ggt	624
Gly Glu Val His Ile Val Glu Ala Asp Arg Asn Ser Thr Gly Ala Gly	
195 200 205	
gac gac ttg tgg tgg gcg cac acc ggt ggc ggt ggc ggc aac ttc ggg	672
His Asp Leu Trp Trp Ala His Thr Gly Gly Gly Gly Gly Asn Phe Gly	
210 215 220	
atc gtc acc agg ttt tgg ctc cga acg ccg gac gtg gtc agc acc gac	720
Ile Val Thr Arg Phe Trp Leu Arg Thr Pro Asp Val Val Ser Thr Asp	
225 230 235 240	
gcc gca gag ctc ctg cca cgg ccg ccc gcg aca gtg ctg ctc cga tcg	768
Ala Ala Glu Leu Leu Pro Arg Pro Pro Ala Thr Val Leu Leu Arg Ser	
245 250 255	
ttc cac tgg ccg tgg cac gaa ctg aca gag cag tca ttc gcc gtg ctc	816
Phe His Trp Pro Trp His Glu Leu Thr Glu Gln Ser Phe Ala Val Leu	
260 265 270	

cta cag aac ttc ggc aat tgg tac gag cag cac agc gcg cct gaa tcc	864
Leu Gln Asn Phe Gly Asn Trp Tyr Glu Gln His Ser Ala Pro Glu Ser	
275 280 285	
acg caa ctc ggg ttg ttc agc acg ctc gtc tgc gca cac cgg caa gct	912
Thr Gln Leu Gly Leu Phe Ser Thr Leu Val Cys Ala His Arg Gln Ala	
290 295 300	
ggc tac gtc acg ctg aac gtt cac ctg gac ggc acg gat ccg aac gcg	960
Gly Tyr Val Thr Leu Asn Val His Leu Asp Gly Thr Asp Pro Asn Ala	
305 310 315 320	
gaa cgc acc ctg gcc gaa cac ctg tcg gcg atc aac gcc cag gtc ggc	1008
Glu Arg Thr Leu Ala Glu His Leu Ser Ala Ile Asn Ala Gln Val Gly	
325 330 335	
gtg act cca gcc gaa ggg ctg cgg gaa acc ctg ccg tgg ttg cga tcg	1056
Val Thr Pro Ala Glu Gly Leu Arg Glu Thr Leu Pro Trp Leu Arg Ser	
340 345 350	
acc cag gtg gcc ggg gcg atc gcc gaa ggc ggc gaa ccg ggc atg caa	1104
Thr Gln Val Ala Gly Ala Ile Ala Glu Gly Gly Glu Pro Gly Met Gln	
355 360 365	
cgg acc aag gtc aaa gcc gcc tac ttg cgc acc ggg ctg tcc gaa gct	1152
Arg Thr Lys Val Lys Ala Ala Tyr Leu Arg Thr Gly Leu Ser Glu Ala	
370 375 380	
caa cta gcc acg gtt tac cgg cgg ctg acc gtc tac gga tac gac aac	1200
Gln Leu Ala Thr Val Tyr Arg Arg Leu Thr Val Tyr Gly Tyr Asp Asn	
385 390 395 400	
cct gcg gcg gcg ctg ttg ctg ctc ggt tac ggc ggt atg gcg aat gcc	1248
Pro Ala Ala Ala Leu Leu Leu Gly Tyr Gly Gly Met Ala Asn Ala	
405 410 415	
gtg gct ccg tcg gcc acc gca ctc gct cag cgc gac tcg gtt ctc aaa	1296
Val Ala Pro Ser Ala Thr Ala Leu Ala Gln Arg Asp Ser Val Leu Lys	
420 425 430	
ggc ctg ttc gtc acg aac tgg tcg gag ccc gcc gag gac gag cgg cat	1344
Ala Leu Phe Val Thr Asn Trp Ser Glu Pro Ala Glu Asp Glu Arg His	
435 440 445	
ctg acc tgg att cgc ggt ttc tac cgc gag atg tac gcc gaa acc ggc	1392
Leu Thr Trp Ile Arg Gly Phe Tyr Arg Glu Met Tyr Ala Glu Thr Gly	
450 455 460	

gga gtt ccg gtg cca ggt acc cgt gtc gac ggc tcc tac atc aac tac 1440  
 Gly Val Pro Val Pro Gly Thr Arg Val Asp Gly Ser Tyr Ile Asn Tyr  
 465 470 475 480

ccg gac acc gac ttg gcc gat cca ttg tgg aac acc tcc ggt gtt gcc 1488  
 Pro Asp Thr Asp Leu Ala Asp Pro Leu Trp Asn Thr Ser Gly Val Ala  
 485 490 495

tgg cac gac ctg tac tac aaa gac aac tac ccg cgg ctg cag cgg gcc 1536  
 Trp His Asp Leu Tyr Tyr Lys Asp Asn Tyr Pro Arg Leu Gln Arg Ala  
 500 505 510

aaa gcg cgg tgg gat ccg cag aac atc ttc cag cac ggc ctg tcg atc 1584  
 Lys Ala Arg Trp Asp Pro Gln Asn Ile Phe Gln His Gly Leu Ser Ile  
 515 520 525

aaa ccg ccg gca cgg ctt tca ccc ggt cag cca tga 1620  
 Lys Pro Pro Ala Arg Leu Ser Pro Gly Gln Pro  
 530 535

<210> 16  
 <211> 539  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 16  
 Met Ile Ser Ala Ala Gly Glu Gln Ser Gly Pro Val Arg Lys Gly Gly  
 1 5 10 15

Ala Val Pro Glu Phe His Asp Pro Ala Pro Met Asn Arg Arg Thr Pro  
 20 25 30

Gly Thr Glu Ile Thr Val Glu Pro Asp Asp Pro Arg Tyr Pro Asp Leu  
 35 40 45

Val Val Gly His Asn Pro Arg Phe Thr Gly Lys Pro Glu Arg Ile His  
 50 55 60

Ile Ala Ser Ser Ala Glu Asp Val Val His Ala Val Ala Asp Ala Val  
 65 70 75 80

Arg Thr Gly Arg Arg Val Gly Val Arg Ser Gly Gly His Cys Phe Glu  
 85 90 95

Asn Leu Val Ala Asp Pro Ala Ile Arg Val Leu Val Asp Leu Ser Glu  
 100 105 110

Leu Asn Arg Val Tyr Tyr Asp Ser Thr Arg Gly Ala Phe Ala Ile Glu  
 115 120 125  
 Ala Gly Ala Ala Leu Gly Gln Val Tyr Arg Thr Leu Phe Lys Asn Trp  
 130 135 140  
 Gly Val Thr Ile Pro Thr Gly Ala Cys Pro Gly Val Gly Ala Gly Gly  
 145 150 155 160  
 His Ile Leu Gly Gly Gly Tyr Gly Pro Leu Ser Arg Arg Phe Gly Ser  
 165 170 175  
 Val Val Asp Tyr Leu Gln Gly Val Glu Val Val Val Val Asp Gln Ala  
 180 185 190  
 Gly Glu Val His Ile Val Glu Ala Asp Arg Asn Ser Thr Gly Ala Gly  
 195 200 205  
 His Asp Leu Trp Trp Ala His Thr Gly Gly Gly Gly Gly Asn Phe Gly  
 210 215 220  
 Ile Val Thr Arg Phe Trp Leu Arg Thr Pro Asp Val Val Ser Thr Asp  
 225 230 235 240  
 Ala Ala Glu Leu Leu Pro Arg Pro Pro Ala Thr Val Leu Leu Arg Ser  
 245 250 255  
 Phe His Trp Pro Trp His Glu Leu Thr Glu Gln Ser Phe Ala Val Leu  
 260 265 270  
 Leu Gln Asn Phe Gly Asn Trp Tyr Glu Gln His Ser Ala Pro Glu Ser  
 275 280 285  
 Thr Gln Leu Gly Leu Phe Ser Thr Leu Val Cys Ala His Arg Gln Ala  
 290 295 300  
 Gly Tyr Val Thr Leu Asn Val His Leu Asp Gly Thr Asp Pro Asn Ala  
 305 310 315 320  
 Glu Arg Thr Leu Ala Glu His Leu Ser Ala Ile Asn Ala Gln Val Gly  
 325 330 335  
 Val Thr Pro Ala Glu Gly Leu Arg Glu Thr Leu Pro Trp Leu Arg Ser  
 340 345 350  
 Thr Gln Val Ala Gly Ala Ile Ala Glu Gly Gly Glu Pro Gly Met Gln  
 355 360 365

Arg Thr Lys Val Lys Ala Ala Tyr Leu Arg Thr Gly Leu Ser Glu Ala  
370 375 380

Gln Leu Ala Thr Val Tyr Arg Arg Leu Thr Val Tyr Gly Tyr Asp Asn  
385 390 395 400

Pro Ala Ala Ala Leu Leu Leu Leu Gly Tyr Gly Gly Met Ala Asn Ala  
405 410 415

Val Ala Pro Ser Ala Thr Ala Leu Ala Gln Arg Asp Ser Val Leu Lys  
420 425 430

Ala Leu Phe Val Thr Asn Trp Ser Glu Pro Ala Glu Asp Glu Arg His  
435 440 445

Leu Thr Trp Ile Arg Gly Phe Tyr Arg Glu Met Tyr Ala Glu Thr Gly  
450 455 460

Gly Val Pro Val Pro Gly Thr Arg Val Asp Gly Ser Tyr Ile Asn Tyr  
465 470 475 480

Pro Asp Thr Asp Leu Ala Asp Pro Leu Trp Asn Thr Ser Gly Val Ala  
485 490 495

Trp His Asp Leu Tyr Tyr Lys Asp Asn Tyr Pro Arg Leu Gln Arg Ala  
500 505 510

Lys Ala Arg Trp Asp Pro Gln Asn Ile Phe Gln His Gly Leu Ser Ile  
515 520 525

Lys Pro Pro Ala Arg Leu Ser Pro Gly Gln Pro  
530 535

<210> 17

<211> 1194

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1191)

<223> ORF6; Methyltransferase

<400> 17

atg tcc aca acg cac gag atc gaa acc gtg gaa cgc atc atc ctc gcc 48

Met	Ser	Thr	Thr	His	Glu	Ile	Glu	Thr	Val	Glu	Arg	Ile	Leu	Ala	
1				5				10					15		
gcc gga tcc agt gcg gcg agc ctg gcc gac ctg acc acc gaa ctc gga															96
Ala Gly Ser Ser Ala Ala Ser Leu Ala Asp Leu Thr Thr Glu Leu Gly															
				20				25					30		
ctc gcc agg atc gca ccc gtg ctg atc gac gag atc ctc ttc cgc gcg															144
Leu Ala Arg Ile Ala Pro Val Leu Ile Asp Glu Ile Leu Phe Arg Ala															
				35				40					45		
gaa ccg gcc ccc gac atc gaa cgg acc gag gtc gcg gtc cag atc acc															192
Glu Pro Ala Pro Asp Ile Glu Arg Thr Glu Val Ala Val Gln Ile Thr															
				50				55					60		
cac cga ggc gag acc gtt gac ttc gtc ctg acg cta cag tcc ggt gag															240
His Arg Gly Glu Thr Val Asp Phe Val Leu Thr Leu Gln Ser Gly Glu															
				65				70					75		80
ctg atc aag gcc gag caa cga ccg gtc gga gac gtc ccg ctg cgg atc															288
Leu Ile Lys Ala Glu Gln Arg Pro Val Gly Asp Val Pro Leu Arg Ile															
				85				90					95		
ggt tac gag ctc acc gat ctc atc gcc gag ttg ttc ggc cca gga gct															336
Gly Tyr Glu Leu Thr Asp Leu Ile Ala Glu Leu Phe Gly Pro Gly Ala															
				100				105					110		
ccc agg gcc gtc ggc gcc cgg agc acc aac ttc ctc cga acc acc aca															384
Pro Arg Ala Val Gly Ala Arg Ser Thr Asn Phe Leu Arg Thr Thr Thr															
				115				120					125		
tcc ggt tcg ata ccc ggt ccg tcg gaa ctg tcc gat ggc ttc cag gcc															432
Ser Gly Ser Ile Pro Gly Pro Ser Glu Leu Ser Asp Gly Phe Gln Ala															
				130				135					140		
atc tcc gca gtg gtc gcc ggc tgc ggg cac cga cgt ccc gac ctc aac															480
Ile Ser Ala Val Val Ala Gly Cys Gly His Arg Arg Pro Asp Leu Asn															
				145				150					155		160
ttg ctc gcc tcc cac tac cgc acg gac aag tgg ggc ggc ctg cac tgg															528
Leu Leu Ala Ser His Tyr Arg Thr Asp Lys Trp Gly Gly Leu His Trp															
				165				170					175		
ttc acc ccg cta tac gag cga cac ctc ggc gag ttc cgt gat cgc ccg															576
Phe Thr Pro Leu Tyr Glu Arg His Leu Gly Glu Phe Arg Asp Arg Pro															
				180				185					190		
gtg cgc atc ctg gag atc ggt gtc ggt ggc tac aac ttc gac ggt ggc															624



Val Arg Ile Leu Glu Ile Gly Val Gly Gly Tyr Asn Phe Asp Gly Gly	
195	200 205
ggc ggc gaa tcc ctg aag atg tgg aag cgc tac ttc cac cgc ggc ctc	672
Gly Gly Glu Ser Leu Lys Met Trp Lys Arg Tyr Phe His Arg Gly Leu	
210 215 220	
gtg ttc ggg atg gac gtt ttc gac aag tcc ttc ctc gac cag cag agg	720
Val Phe Gly Met Asp Val Phe Asp Lys Ser Phe Leu Asp Gln Gln Arg	
225 230 235 240	
ctc tgc acc gtc cgc gcc gac cag agc aag ccc gag gag ctg gcc gcc	768
Leu Cys Thr Val Arg Ala Asp Gln Ser Lys Pro Glu Glu Leu Ala Ala	
245 250 255	
gtt gac gac aag tac gga ccg ttc gac atc atc atc gac gat ggc agc	816
Val Asp Asp Lys Tyr Gly Pro Phe Asp Ile Ile Ile Asp Asp Gly Ser	
260 265 270	
cac atc aac gga cac gtg cgc aca tcc ctg gaa acg ctg ttc ccc cgg	864
His Ile Asn Gly His Val Arg Thr Ser Leu Glu Thr Leu Phe Pro Arg	
275 280 285	
ttg cgc agc ggt ggc gta tac gtg atc gag gat ctg tgg acg acc tat	912
Leu Arg Ser Gly Gly Val Tyr Val Ile Glu Asp Leu Trp Thr Thr Tyr	
290 295 300	
gct ccc gga ttc ggc ggg cag gcg cag tgc ccg gcc gca ccc ggc acc	960
Ala Pro Gly Phe Gly Gly Gln Ala Gln Cys Pro Ala Ala Pro Gly Thr	
305 310 315 320	
acg gtc agc ctg ctc aag aac ctg ttg gaa ggc gtt cag cac gag gag	1008
Thr Val Ser Leu Leu Lys Asn Leu Leu Glu Gly Val Gln His Glu Glu	
325 330 335	
cag ccg cat gcg ggc tcg tac gag ccg agc tac ctg gaa cgc aat ttg	1056
Gln Pro His Ala Gly Ser Tyr Glu Pro Ser Tyr Leu Glu Arg Asn Leu	
340 345 350	
gtc ggc ctc cac acc tac cac aac atc gcg ttc ctg gag aaa ggc gtc	1104
Val Gly Leu His Thr Tyr His Asn Ile Ala Phe Leu Glu Lys Gly Val	
355 360 365	
aac gcc gaa ggc ggc gtt cct gct tgg gtg cca agg agt ctg gac gac	1152
Asn Ala Glu Gly Gly Val Pro Ala Trp Val Pro Arg Ser Leu Asp Asp	
370 375 380	
ata ttg cac ctg gcc gac gtg aac agc gcg gag gac gag tga	1194

Ile Leu His Leu Asp Val Asn Ser Ala Glu Asp  
385 390 395

<210> 18

<211> 397

<212> PRT

<213> Saccharopolyspora spinosa

<400> 18

Met Ser Thr Thr His Glu Ile Glu Thr Val Glu Arg Ile Ile Leu Ala  
1 5 10 15

Ala Gly Ser Ser Ala Ala Ser Leu Ala Asp Leu Thr Thr Glu Leu Gly  
20 25 30

Leu Ala Arg Ile Ala Pro Val Leu Ile Asp Glu Ile Leu Phe Arg Ala  
35 40 45

Glu Pro Ala Pro Asp Ile Glu Arg Thr Glu Val Ala Val Gln Ile Thr  
50 55 60

His Arg Gly Glu Thr Val Asp Phe Val Leu Thr Leu Gln Ser Gly Glu  
65 70 75 80

Leu Ile Lys Ala Glu Gln Arg Pro Val Gly Asp Val Pro Leu Arg Ile  
85 90 95

Gly Tyr Glu Leu Thr Asp Leu Ile Ala Glu Leu Phe Gly Pro Gly Ala  
100 105 110

Pro Arg Ala Val Gly Ala Arg Ser Thr Asn Phe Leu Arg Thr Thr Thr  
115 120 125

Ser Gly Ser Ile Pro Gly Pro Ser Glu Leu Ser Asp Gly Phe Gln Ala  
130 135 140

Ile Ser Ala Val Val Ala Gly Cys Gly His Arg Arg Pro Asp Leu Asn  
145 150 155 160

Leu Leu Ala Ser His Tyr Arg Thr Asp Lys Trp Gly Gly Leu His Trp  
165 170 175

Phe Thr Pro Leu Tyr Glu Arg His Leu Gly Glu Phe Arg Asp Arg Pro  
180 185 190

Val Arg Ile Leu Glu Ile Gly Val Gly Gly Tyr Asn Phe Asp Gly Gly  
195 200 205

Gly Gly Glu Ser Leu Lys Met Trp Lys Arg Tyr Phe His Arg Gly Leu  
210 215 220

Val Phe Gly Met Asp Val Phe Asp Lys Ser Phe Leu Asp Gln Gln Arg  
225 230 235 240

Leu Cys Thr Val Arg Ala Asp Gln Ser Lys Pro Glu Glu Leu Ala Ala  
245 250 255

Val Asp Asp Lys Tyr Gly Pro Phe Asp Ile Ile Ile Asp Asp Gly Ser  
260 265 270

His Ile Asn Gly His Val Arg Thr Ser Leu Glu Thr Leu Phe Pro Arg  
275 280 285

Leu Arg Ser Gly Gly Val Tyr Val Ile Glu Asp Leu Trp Thr Thr Tyr  
290 295 300

Ala Pro Gly Phe Gly Gly Gln Ala Gln Cys Pro Ala Ala Pro Gly Thr  
305 310 315 320

Thr Val Ser Leu Leu Lys Asn Leu Leu Glu Gly Val Gln His Glu Glu  
325 330 335

Gln Pro His Ala Gly Ser Tyr Glu Pro Ser Tyr Leu Glu Arg Asn Leu  
340 345 350

Val Gly Leu His Thr Tyr His Asn Ile Ala Phe Leu Glu Lys Gly Val  
355 360 365

Asn Ala Glu Gly Gly Val Pro Ala Trp Val Pro Arg Ser Leu Asp Asp  
370 375 380

Ile Leu His Leu Ala Asp Val Asn Ser Ala Glu Asp Glu  
385 390 395

<210> 19

<211> 900

<212> DNA

<213> *Saccharopolyspora spinosa*

<220>

<221> CDS

<222> (1)..(897)

<223> ORF7; O-Methyltransferase

<400> 19

gtg aac agc aga ggg gcg aac aca cag gca ttt ccg acc gcg gat cag 48  
Val Asn Ser Arg Gly Ala Asn Thr Gln Ala Phe Pro Thr Ala Asp Gln  
1 5 10 15

gtg gag tcc atc ttc gat gcg ttg gcg cac ggg cgt ccc ctg cac cac 96  
Val Glu Ser Ile Phe Asp Ala Leu Ala His Gly Arg Pro Leu His His  
20 25 30

ggt tac tgg gcg ggc ggg tat cgg gag gat gcc ggt gcc aca ccg tgg 144  
Gly Tyr Trp Ala Gly Gly Tyr Arg Glu Asp Ala Gly Ala Thr Pro Trp  
35 40 45

tcg gat gct gcc gac caa ctg acc gac ctg ttc atc gac aag gcc gcg 192  
Ser Asp Ala Ala Asp Gln Leu Thr Asp Leu Phe Ile Asp Lys Ala Ala  
50 55 60

ctc cgt ccc gga gcg cac ctg ttc gac ctg ggc tgc ggc aat ggg cag 240  
Leu Arg Pro Gly Ala His Leu Phe Asp Leu Gly Cys Gly Asn Gly Gln  
65 70 75 80

ccc gta gtc cgt gcg gca tgc gcc agc ggc gtt cga gtc acc gga atc 288  
Pro Val Val Arg Ala Ala Cys Ala Ser Gly Val Arg Val Thr Gly Ile  
85 90 95

acc gtg aac gcc cag cat ctc gcc gcc gcc acc agg ctc gcc aac gag 336  
Thr Val Asn Ala Gln His Leu Ala Ala Ala Thr Arg Leu Ala Asn Glu  
100 105 110

acc gga ctg gcc ggc agt ctt gag ttc gat cta gtc gac ggc gcc cag 384  
Thr Gly Leu Ala Gly Ser Leu Glu Phe Asp Leu Val Asp Gly Ala Gln  
115 120 125

ctg ccc tac ccg gac ggt ttc ttt cag gcc gca tgg gcg atg cag tcc 432  
Leu Pro Tyr Pro Asp Gly Phe Phe Gln Ala Ala Trp Ala Met Gln Ser  
130 135 140

gtc gtg cag atc gtg gac cag gcc gcc gcg atc cgc gag gtc cac cga 480  
Val Val Gln Ile Val Asp Gln Ala Ala Ala Ile Arg Glu Val His Arg  
145 150 155 160

atc ctg gaa ccc ggc ggc cgg ttc gtc ctc gga gac atc atc act cgg 528  
Ile Leu Glu Pro Gly Gly Arg Phe Val Leu Gly Asp Ile Ile Thr Arg  
165 170 175

gtt cga ctc ccg gaa gag tac gcg gcg gtt tgg acg ggc acg acc gcc 576  
Val Arg Leu Pro Glu Glu Tyr Ala Ala Val Trp Thr Gly Thr Thr Ala

180	185	190	
cat acc ttg aac agc ttc acg gcg ctg gtc agc gaa gcc ggg ttc gag			624
His Thr Leu Asn Ser Phe Thr Ala Leu Val Ser Glu Ala Gly Phe Glu			
195	200	205	
att ctc gaa gtc acc gac ctc acg gca cag acc agg tgc atg gtc tcc			672
Ile Leu Glu Val Thr Asp Leu Thr Ala Gln Thr Arg Cys Met Val Ser			
210	215	220	
tgg tac gtc gac gag ttg ctc cgg aaa ctc gat gag ctc gcc ggc gtc			720
Trp Tyr Val Asp Glu Leu Leu Arg Lys Leu Asp Glu Leu Ala Gly Val			
225	230	235	240
gag cct gcg gct gtc ggc acc tac cag caa cgc tac ttg gga gac atc			768
Glu Pro Ala Ala Val Gly Thr Tyr Gln Gln Arg Tyr Leu Gly Asp Ile			
245	250	255	
gcg gcg aag cac gga ccg gga cca gca cag ctg atc gcc gcg gtt gcg			816
Ala Ala Lys His Gly Pro Gly Pro Ala Gln Leu Ile Ala Ala Val Ala			
260	265	270	
gaa tac cgg aaa cat ccg gat tac gcc aga aac gag gaa agc atg ggt			864
Glu Tyr Arg Lys His Pro Asp Tyr Ala Arg Asn Glu Glu Ser Met Gly			
275	280	285	
ttc atg ctc ctg cag gct cga aag aag cag tcc tga			900
Phe Met Leu Leu Gln Ala Arg Lys Lys Gln Ser			
290	295		

<210> 20

<211> 299

<212> PRT

<213> Saccharopolyspora spinosa

<400> 20

Val Asn Ser Arg Gly Ala Asn Thr Gln Ala Phe Pro Thr Ala Asp Gln
1 5 10 15

Val Glu Ser Ile Phe Asp Ala Leu Ala His Gly Arg Pro Leu His His
20 25 30

Gly Tyr Trp Ala Gly Gly Tyr Arg Glu Asp Ala Gly Ala Thr Pro Trp
35 40 45

Ser Asp Ala Ala Asp Gln Leu Thr Asp Leu Phe Ile Asp Lys Ala Ala
50 55 60

Leu Arg Pro Gly Ala His Leu Phe Asp Leu Gly Cys Gly Asn Gly Gln  
 65 70 75 80

Pro Val Val Arg Ala Ala Cys Ala Ser Gly Val Arg Val Thr Gly Ile  
 85 90 95

Thr Val Asn Ala Gln His Leu Ala Ala Ala Thr Arg Leu Ala Asn Glu  
 100 105 110

Thr Gly Leu Ala Gly Ser Leu Glu Phe Asp Leu Val Asp Gly Ala Gln  
 115 120 125

Leu Pro Tyr Pro Asp Gly Phe Phe Gln Ala Ala Trp Ala Met Gln Ser  
 130 135 140

Val Val Gln Ile Val Asp Gln Ala Ala Ala Ile Arg Glu Val His Arg  
 145 150 155 160

Ile Leu Glu Pro Gly Gly Arg Phe Val Leu Gly Asp Ile Ile Thr Arg  
 165 170 175

Val Arg Leu Pro Glu Glu Tyr Ala Ala Val Trp Thr Gly Thr Thr Ala  
 180 185 190

His Thr Leu Asn Ser Phe Thr Ala Leu Val Ser Glu Ala Gly Phe Glu  
 195 200 205

Ile Leu Glu Val Thr Asp Leu Thr Ala Gln Thr Arg Cys Met Val Ser  
 210 215 220

Trp Tyr Val Asp Glu Leu Leu Arg Lys Leu Asp Glu Leu Ala Gly Val  
 225 230 235 240

Glu Pro Ala Ala Val Gly Thr Tyr Gln Gln Arg Tyr Leu Gly Asp Ile  
 245 250 255

Ala Ala Lys His Gly Pro Gly Pro Ala Gln Leu Ile Ala Ala Val Ala  
 260 265 270

Glu Tyr Arg Lys His Pro Asp Tyr Ala Arg Asn Glu Glu Ser Met Gly  
 275 280 285

Phe Met Leu Leu Gln Ala Arg Lys Lys Gln Ser  
 290 295

<210> 21  
 <211> 1167  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(1164)  
 <223> ORF8; Cyclisierungsenzym

<400> 21  
 atg gcc tcc gag cac gcc agc ctg gtc ggc gac gat ctg cgg gca ccc 48  
 Met Ala Ser Glu His Ala Ser Leu Val Gly Asp Asp Leu Arg Ala Pro  
 1 5 10 15  
 gcg gat gat ccc ttc tac cga ccg ccg acg ccg cta ccg ccg ggt gtc 96  
 Ala Asp Asp Pro Phe Tyr Arg Pro Pro Thr Pro Leu Pro Pro Gly Val  
 20 25 30  
 ccg ggc acg ctc ctc agg gcc ccg ccc gtc tcg gca ctg cgc ggc acg 144  
 Pro Gly Thr Leu Leu Arg Ala Arg Pro Val Ser Ala Leu Arg Gly Thr  
 35 40 45  
 ggc gaa ccc gtc gca gcc aag gcc tgg caa atc ctc tac cgg tcc aac 192  
 Gly Glu Pro Val Ala Ala Lys Ala Trp Gln Ile Leu Tyr Arg Ser Asn  
 50 55 60  
 tcc gcc ctt ggc atg ccg aac gcc gtc tcc ggc acc gtt ctg gtg ccg 240  
 Ser Ala Leu Gly Met Pro Asn Ala Val Ser Gly Thr Val Leu Val Pro  
 65 70 75 80  
 aac atc ccg tgg ccg cgc gaa gat cgc ccc atc atc act ttc gca gtg 288  
 Asn Ile Pro Trp Pro Arg Glu Asp Arg Pro Ile Ile Thr Phe Ala Val  
 85 90 95  
 ggc acc cac ggc ctc ggt agc caa gtt gcc ccg tcg tac ctg ctt cga 336  
 Gly Thr His Gly Leu Gly Ser Gln Val Ala Pro Ser Tyr Leu Leu Arg  
 100 105 110  
 acc gga acc gag ccg gag acc gag ctg atc gcc gtg gcc ctc gac cgc 384  
 Thr Gly Thr Glu Pro Glu Thr Glu Leu Ile Ala Val Ala Leu Asp Arg  
 115 120 125  
 ggg tgg gcc gtg gtc atc acc gac tac gaa ggc ctc ggt act cct gga 432  
 Gly Trp Ala Val Val Ile Thr Asp Tyr Glu Gly Leu Gly Thr Pro Gly  
 130 135 140  
 acc cac acc tac acc gtc ggc agg gcg cag gga cac gcc atg ctc gat 480

Thr His Thr Tyr Thr Val Gly Arg Ala Gln Gly His Ala Met Leu Asp 145 150 155 160  
 gcc gcc cgc gct gcg caa cgg cta ccg ggc tcc ggc ctg acg acc gac 528  
 Ala Ala Arg Ala Ala Gln Arg Leu Pro Gly Ser Gly Leu Thr Thr Asp  
 165 170 175  
 tgc ccg gtc ggc atc tgg ggc tat gcg cag ggt ggg caa gcg tcg gcc 576  
 Cys Pro Val Gly Ile Trp Gly Tyr Ala Gln Gly Gly Gln Ala Ser Ala  
 180 185 190  
 ttc gcc ggc gaa ctg cac ccc acc tac gca cct gaa ctg cga atc cgc 624  
 Phe Ala Gly Glu Leu His Pro Thr Tyr Ala Pro Glu Leu Arg Ile Arg  
 195 200 205  
 gct gcg gcc gca ggt gcg gtg ccg atc gat ctg ctg gac atc atc cac 672  
 Ala Ala Ala Gly Ala Val Pro Ile Asp Leu Leu Asp Ile Ile His  
 210 215 220  
 cga aat gac ggg gtg ttc acc ggg ccg gtg ctg gcc ggc ctg gtc ggg 720  
 Arg Asn Asp Gly Val Phe Thr Gly Pro Val Leu Ala Gly Leu Val Gly  
 225 230 235 240  
 cat gcc gct gcc tac ccc gat ctg cca ttc gac gag ctt ctc acc gaa 768  
 His Ala Ala Ala Tyr Pro Asp Leu Pro Phe Asp Glu Leu Leu Thr Glu  
 245 250 255  
 gcg ggt cgt acc gcc gtt gat caa gtg cgc gag ctc ggt gca ccg gag 816  
 Ala Gly Arg Thr Ala Val Asp Gln Val Arg Glu Leu Gly Ala Pro Glu  
 260 265 270  
 ctc gtc acc cgc ttc ctc ggc cgc gag ctg agc gac ttc ctc gac act 864  
 Leu Val Thr Arg Phe Leu Gly Arg Glu Leu Ser Asp Phe Leu Asp Thr  
 275 280 285  
 tcc ggc ctt ttc gag caa cct cga tgg cgc gca cga ctg gcc gaa agc 912  
 Ser Gly Leu Phe Glu Gln Pro Arg Trp Arg Ala Arg Leu Ala Glu Ser  
 290 295 300  
 gtc gca ggt agg aac ggt ggc ccg gtg gtc ccc acg ctc gtc tac cac 960  
 Val Ala Gly Arg Asn Gly Gly Pro Val Val Pro Thr Leu Val Tyr His  
 305 310 315 320  
 agc acg gac gac gag atc gtt ccg ttc gca ttc ggc gag cga ctc cgg 1008  
 Ser Thr Asp Asp Glu Ile Val Pro Phe Ala Phe Gly Glu Arg Leu Arg  
 325 330 335  
 gac agc tac cgc gcg gcg ggt acg cca gtg cgg tgg cat ccg ctc tcc 1056



Asp Ser Tyr Arg Ala Ala Gly Thr Pro Val Arg Trp His Pro Leu Ser  
 340 345 350

gga ttg gct cac ttt ccc gcc gcc ctg gcc agc tcg cga gtg gtc gtc 1104  
 Gly Leu Ala His Phe Pro Ala Ala Leu Ala Ser Ser Arg Val Val Val  
 355 360 365

tcg tgg ttc gac gag cac ttc tcc gag ccg tcc gcg atc agc ggt ccg 1152  
 Ser Trp Phe Asp Glu His Phe Ser Glu Pro Ser Ala Ile Ser Gly Pro  
 370 375 380

cga gat gcc agg tga 1167  
 Arg Asp Ala Arg  
 385

<210> 22

<211> 388

<212> PRT

<213> Saccharopolyspora spinosa

<400> 22

Met Ala Ser Glu His Ala Ser Leu Val Gly Asp Asp Leu Arg Ala Pro  
 1 5 10 15

Ala Asp Asp Pro Phe Tyr Arg Pro Pro Thr Pro Leu Pro Pro Gly Val  
 20 25 30

Pro Gly Thr Leu Leu Arg Ala Arg Pro Val Ser Ala Leu Arg Gly Thr  
 35 40 45

Gly Glu Pro Val Ala Ala Lys Ala Trp Gln Ile Leu Tyr Arg Ser Asn  
 50 55 60

Ser Ala Leu Gly Met Pro Asn Ala Val Ser Gly Thr Val Leu Val Pro  
 65 70 75 80

Asn Ile Pro Trp Pro Arg Glu Asp Arg Pro Ile Ile Thr Phe Ala Val  
 85 90 95

Gly Thr His Gly Leu Gly Ser Gln Val Ala Pro Ser Tyr Leu Leu Arg  
 100 105 110

Thr Gly Thr Glu Pro Glu Thr Glu Leu Ile Ala Val Ala Leu Asp Arg  
 115 120 125

Gly Trp Ala Val Val Ile Thr Asp Tyr Glu Gly Leu Gly Thr Pro Gly  
 130 135 140

Thr His Thr Tyr Thr Val Gly Arg Ala Gln Gly His Ala Met Leu Asp  
 145 150 155 160

Ala Ala Arg Ala Ala Gln Arg Leu Pro Gly Ser Gly Leu Thr Thr Asp  
 165 170 175

Cys Pro Val Gly Ile Trp Gly Tyr Ala Gln Gly Gly Gln Ala Ser Ala  
 180 185 190

Phe Ala Gly Glu Leu His Pro Thr Tyr Ala Pro Glu Leu Arg Ile Arg  
 195 200 205

Ala Ala Ala Ala Gly Ala Val Pro Ile Asp Leu Leu Asp Ile Ile His  
 210 215 220

Arg Asn Asp Gly Val Phe Thr Gly Pro Val Leu Ala Gly Leu Val Gly  
 225 230 235 240

His Ala Ala Ala Tyr Pro Asp Leu Pro Phe Asp Glu Leu Leu Thr Glu  
 245 250 255

Ala Gly Arg Thr Ala Val Asp Gln Val Arg Glu Leu Gly Ala Pro Glu  
 260 265 270

Leu Val Thr Arg Phe Leu Gly Arg Glu Leu Ser Asp Phe Leu Asp Thr  
 275 280 285

Ser Gly Leu Phe Glu Gln Pro Arg Trp Arg Ala Arg Leu Ala Glu Ser  
 290 295 300

Val Ala Gly Arg Asn Gly Gly Pro Val Val Pro Thr Leu Val Tyr His  
 305 310 315 320

Ser Thr Asp Asp Glu Ile Val Pro Phe Ala Phe Gly Glu Arg Leu Arg  
 325 330 335

Asp Ser Tyr Arg Ala Ala Gly Thr Pro Val Arg Trp His Pro Leu Ser  
 340 345 350

Gly Leu Ala His Phe Pro Ala Ala Leu Ala Ser Ser Arg Val Val Val  
 355 360 365

Ser Trp Phe Asp Glu His Phe Ser Glu Pro Ser Ala Ile Ser Gly Pro  
 370 375 380

Arg Asp Ala Arg  
 385

<210> 23  
 <211> 1011  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(1008)  
 <223> ORF9; 2,3-Reduktase

<400> 23  
 atg acc agc tgc atg cga aag ccg gtg cgc atc ggt gtg ctc ggg tgc 48  
 Met Thr Ser Ser Met Arg Lys Pro Val Arg Ile Gly Val Leu Gly Cys  
 1 5 10 15  
 gct tcc ttc gcg tgg cga cgg atg ctg ccc gcg atg tgc gac gtg gcc 96  
 Ala Ser Phe Ala Trp Arg Arg Met Leu Pro Ala Met Cys Asp Val Ala  
 20 25 30  
 gaa aca gag gtg gtg gcg gtg gcg agc cgt gat ccg gcg aaa gcc gaa 144  
 Glu Thr Glu Val Val Ala Val Ala Ser Arg Asp Pro Ala Lys Ala Glu  
 35 40 45  
 cgg ttc gca gcg cga ttc gaa tgc gag gcg gtg ctg ggt tac cag cgg 192  
 Arg Phe Ala Ala Arg Phe Glu Cys Glu Ala Val Leu Gly Tyr Gln Arg  
 50 55 60  
 ctc ctg gag ccg ccg gac atc gat gcc gtc tac gtg ccg ttg ccg cct 240  
 Leu Leu Glu Arg Pro Asp Ile Asp Ala Val Tyr Val Pro Leu Pro Pro  
 65 70 75 80  
 ggc atg cat gca gag tgg atc ggc aag gcg ctt gag gca gac aaa cac 288  
 Gly Met His Ala Glu Trp Ile Gly Lys Ala Leu Glu Ala Asp Lys His  
 85 90 95  
 gtg ctt gcg gag aaa ccg ctg acg acg acg gcg tcc gac acc gct cgc 336  
 Val Leu Ala Glu Lys Pro Leu Thr Thr Thr Ala Ser Asp Thr Ala Arg  
 100 105 110  
 ctg gtc ggg ctg gcc agg agg aag aac ctg ctg ctg ccg gag aat tac 384  
 Leu Val Gly Leu Ala Arg Arg Lys Asn Leu Leu Leu Arg Glu Asn Tyr  
 115 120 125  
 ctg ttc ctc cac cac ggc ccg cac gac gtg gtc cgc gac ctg ctg caa 432  
 Leu Phe Leu His His Gly Arg His Asp Val Val Arg Asp Leu Leu Gln

130	135	140	
tcc ggg gag atc ggt gag ctc cgg gag ttc acc gcc gtg ttc gga att			480
Ser Gly Glu Ile Gly Glu Leu Arg Glu Phe Thr Ala Val Phe Gly Ile			
145	150	155	160
ccg ccg ctt ccc gac acg gac atc cgc tat cgc acc gaa ctc ggt ggc			528
Pro Pro Leu Pro Asp Thr Asp Ile Arg Tyr Arg Thr Glu Leu Gly Gly			
165	170		175
gga gcg ttg ctg gac atc ggt gtc tat ccc gcc cgt gcc gct cgg cac			576
Gly Ala Leu Leu Asp Ile Gly Val Tyr Pro Ala Arg Ala Ala Arg His			
180	185		190
ttt ctc ctc ggt ccg ctc acg gtt ctc ggc gca agc tcg cac gag gcc			624
Phe Leu Leu Gly Pro Leu Thr Val Leu Gly Ala Ser Ser His Glu Ala			
195	200		205
cag gag tcg ggc gtc gac ttg tcg ggc agc gtg ctg ctc caa tcg gaa			672
Gln Glu Ser Gly Val Asp Leu Ser Gly Ser Val Leu Leu Gln Ser Glu			
210	215		220
ggt ggc acc gtt gcc cac ctc gga tac ggt ttc gtg cac cac tac cgc			720
Gly Gly Thr Val Ala His Leu Gly Tyr Gly Phe Val His His Tyr Arg			
225	230	235	240
agc gcg tac gag ctg tgg ggg agt cgt ggg cga atc gtc gtc gac cgg			768
Ser Ala Tyr Glu Leu Trp Gly Ser Arg Gly Arg Ile Val Val Asp Arg			
245	250		255
gcg ttc acg ccg ccc gcc gag tgg cag gcc gtg atc cga atc gag cgg			816
Ala Phe Thr Pro Pro Ala Glu Trp Gln Ala Val Ile Arg Ile Glu Arg			
260	265		270
aag ggc gtt gtc gac gag ttg tcc ttg cca gcg gaa gat cag gtt cgc			864
Lys Gly Val Val Asp Glu Leu Ser Leu Pro Ala Glu Asp Gln Val Arg			
275	280		285
aag gcg gtc acc gcc ttc gca cgc gac atc aga gca ggg aca ggc gtg			912
Lys Ala Val Thr Ala Phe Ala Arg Asp Ile Arg Ala Gly Thr Gly Val			
290	295		300
gac gac cct gcg gtg gcc gga gat tcg ggc gaa tcg atg atc cag cag			960
Asp Asp Pro Ala Val Ala Gly Asp Ser Gly Glu Ser Met Ile Gln Gln			
305	310	315	320
gcc gcg ctg gtg gag gcg atc ggt cag gcc cgt cgg tgc ggg tcc aca			1008
Ala Ala Leu Val Glu Ala Ile Gly Gln Ala Arg Arg Cys Gly Ser Thr			

325

330

335

tag

1011

&lt;210&gt; 24

&lt;211&gt; 336

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 24

Met Thr Ser Ser Met Arg Lys Pro Val Arg Ile Gly Val Leu Gly Cys

1

5

10

15

Ala Ser Phe Ala Trp Arg Arg Met Leu Pro Ala Met Cys Asp Val Ala

20

25

30

Glu Thr Glu Val Val Ala Val Ala Ser Arg Asp Pro Ala Lys Ala Glu

35

40

45

Arg Phe Ala Ala Arg Phe Glu Cys Glu Ala Val Leu Gly Tyr Gln Arg

50

55

60

Leu Leu Glu Arg Pro Asp Ile Asp Ala Val Tyr Val Pro Leu Pro Pro

65

70

75

80

Gly Met His Ala Glu Trp Ile Gly Lys Ala Leu Glu Ala Asp Lys His

85

90

95

Val Leu Ala Glu Lys Pro Leu Thr Thr Thr Ala Ser Asp Thr Ala Arg

100

105

110

Leu Val Gly Leu Ala Arg Arg Lys Asn Leu Leu Leu Arg Glu Asn Tyr

115

120

125

Leu Phe Leu His His Gly Arg His Asp Val Val Arg Asp Leu Leu Gln

130

135

140

Ser Gly Glu Ile Gly Glu Leu Arg Glu Phe Thr Ala Val Phe Gly Ile

145

150

155

160

Pro Pro Leu Pro Asp Thr Asp Ile Arg Tyr Arg Thr Glu Leu Gly Gly

165

170

175

Gly Ala Leu Leu Asp Ile Gly Val Tyr Pro Ala Arg Ala Ala Arg His

180

185

190

Phe Leu Leu Gly Pro Leu Thr Val Leu Gly Ala Ser Ser His Glu Ala

195

200

205

Gln Glu Ser Gly Val Asp Leu Ser Gly Ser Val Leu Leu Gln Ser Glu  
 210 215 220

Gly Gly Thr Val Ala His Leu Gly Tyr Gly Phe Val His His Tyr Arg  
 225 230 235 240

Ser Ala Tyr Glu Leu Trp Gly Ser Arg Gly Arg Ile Val Val Asp Arg  
 245 250 255

Ala Phe Thr Pro Pro Ala Glu Trp Gln Ala Val Ile Arg Ile Glu Arg  
 260 265 270

Lys Gly Val Val Asp Glu Leu Ser Leu Pro Ala Glu Asp Gln Val Arg  
 275 280 285

Lys Ala Val Thr Ala Phe Ala Arg Asp Ile Arg Ala Gly Thr Gly Val  
 290 295 300

Asp Asp Pro Ala Val Ala Gly Asp Ser Gly Glu Ser Met Ile Gln Gln  
 305 310 315 320

Ala Ala Leu Val Glu Ala Ile Gly Gln Ala Arg Arg Cys Gly Ser Thr  
 325 330 335

&lt;210&gt; 25

&lt;211&gt; 1461

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1458)

&lt;223&gt; ORF10; 2,3-Dehydratase

&lt;400&gt; 25

atg agc agt tct gtc gaa gct gag gca agt gct gct gcg ccg ctc ggc 48  
 Met Ser Ser Ser Val Glu Ala Glu Ala Ser Ala Ala Ala Pro Leu Gly  
 1 5 10 15

agc aac aac acg cgg cgg ttc gtc gac tct gcg ctg agc gct tgc aat 96  
 Ser Asn Asn Thr Arg Arg Phe Val Asp Ser Ala Leu Ser Ala Cys Asn  
 20 25 30

ggc atg att ccg acc acg gag ttc cac tgc tgg ctc gcc gat cgg ctg 144

Gly Met Ile Pro Thr Thr Glu Phe His Cys Trp Leu Ala Asp Arg Leu	
35 40 45	
ggc gag aac agc ttc gag acc aat cgc atc ccg ttc gac cgc ctg tcg	192
Gly Glu Asn Ser Phe Glu Thr Asn Arg Ile Pro Phe Asp Arg Leu Ser	
50 55 60	
aaa tgg aaa ttc gat gcc agc acg gag aac ctg gtt cat gcc gac ggt	240
Lys Trp Lys Phe Asp Ala Ser Thr Glu Asn Leu Val His Ala Asp Gly	
65 70 75 80	
agg ttc ttc acg gta gaa ggc ctg cag gtc gag acc aac tat ggc gcg	288
Arg Phe Phe Thr Val Glu Gly Leu Gln Val Glu Thr Asn Tyr Gly Ala	
85 90 95	
gca ccc agc tgg cac cag ccg atc atc aac cag gct gaa gta ggt atc	336
Ala Pro Ser Trp His Gln Pro Ile Ile Asn Gln Ala Glu Val Gly Ile	
100 105 110	
ctc ggc att ctc gtc aag gag atc gac ggc gtg ctg cac tgc ctc atg	384
Leu Gly Ile Leu Val Lys Glu Ile Asp Gly Val Leu His Cys Leu Met	
115 120 125	
tca gca aag atg gaa ccg ggc aac gtc aac gtc ctg cag ctc tcg ccg	432
Ser Ala Lys Met Glu Pro Gly Asn Val Asn Val Leu Gln Leu Ser Pro	
130 135 140	
acg gtt cag gca act cgg agc aac tac acg cag gca cac cgt ggc agc	480
Thr Val Gln Ala Thr Arg Ser Asn Tyr Thr Gln Ala His Arg Gly Ser	
145 150 155 160	
gtt ccg ccc tat gtg gac tac ttc ctc ggg cgg ggc cgc ggc cgc gtg	528
Val Pro Pro Tyr Val Asp Tyr Phe Leu Gly Arg Gly Arg Gly Arg Val	
165 170 175	
ctg gta gac gtg ctc cag tct gaa cag ggg tcc tgg ttc tac cgg aag	576
Leu Val Asp Val Leu Gln Ser Glu Gln Gly Ser Trp Phe Tyr Arg Lys	
180 185 190	
cgc aac cgg aac atg gtg gtg gaa gtc cag gag gaa gtg cca gtc ctg	624
Arg Asn Arg Asn Met Val Val Glu Val Gln Glu Glu Val Pro Val Leu	
195 200 205	
cca gac ttc tgc tgg ttg acg ctc ggc cag gtg ctg gct ctc ctt cgt	672
Pro Asp Phe Cys Trp Leu Thr Leu Gly Gln Val Leu Ala Leu Leu Arg	
210 215 220	
cag gac aac atc gtc aac atg gac acc cgg acg gtg ctg tct tgc atc	720





Ile Leu Ser Glu Glu Gly Gly Arg Phe Tyr Gln Ala Gln Asn Arg Tyr  
 420 425 430

cgg atc atc gag gtg cat gag gac ttc gcg gca cga cct ccc agc gac 1344  
 Arg Ile Ile Glu Val His Glu Asp Phe Ala Ala Arg Pro Pro Ser Asp  
 435 440 445

ttc cgg tgg atg act ttg gga cag ttg ggc gag ctg ctc cgg agc acc 1392  
 Phe Arg Trp Met Thr Leu Gly Gln Leu Gly Glu Leu Leu Arg Ser Thr  
 450 455 460

cac ttc ttg aac atc cag gcg cgc agc ttg gtc gcc tcc ctg cat agc 1440  
 His Phe Leu Asn Ile Gln Ala Arg Ser Leu Val Ala Ser Leu His Ser  
 465 470 475 480

ttg tgg gcg ttg ggg cga tga 1461  
 Leu Trp Ala Leu Gly Arg  
 485

<210> 26  
 <211> 486  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 26  
 Met Ser Ser Ser Val Glu Ala Glu Ala Ser Ala Ala Ala Pro Leu Gly  
 1 5 10 15

Ser Asn Asn Thr Arg Arg Phe Val Asp Ser Ala Leu Ser Ala Cys Asn  
 20 25 30

Gly Met Ile Pro Thr Thr Glu Phe His Cys Trp Leu Ala Asp Arg Leu  
 35 40 45

Gly Glu Asn Ser Phe Glu Thr Asn Arg Ile Pro Phe Asp Arg Leu Ser  
 50 55 60

Lys Trp Lys Phe Asp Ala Ser Thr Glu Asn Leu Val His Ala Asp Gly  
 65 70 75 80

Arg Phe Phe Thr Val Glu Gly Leu Gln Val Glu Thr Asn Tyr Gly Ala  
 85 90 95

Ala Pro Ser Trp His Gln Pro Ile Ile Asn Gln Ala Glu Val Gly Ile  
 100 105 110

Leu Gly Ile Leu Val Lys Glu Ile Asp Gly Val Leu His Cys Leu Met

115

120

125

Ser Ala Lys Met Glu Pro Gly Asn Val Asn Val Leu Gln Leu Ser Pro  
 130 135 140

Thr Val Gln Ala Thr Arg Ser Asn Tyr Thr Gln Ala His Arg Gly Ser  
 145 150 155 160

Val Pro Pro Tyr Val Asp Tyr Phe Leu Gly Arg Gly Arg Gly Arg Val  
 165 170 175

Leu Val Asp Val Leu Gln Ser Glu Gln Gly Ser Trp Phe Tyr Arg Lys  
 180 185 190

Arg Asn Arg Asn Met Val Val Glu Val Gln Glu Glu Val Pro Val Leu  
 195 200 205

Pro Asp Phe Cys Trp Leu Thr Leu Gly Gln Val Leu Ala Leu Leu Arg  
 210 215 220

Gln Asp Asn Ile Val Asn Met Asp Thr Arg Thr Val Leu Ser Cys Ile  
 225 230 235 240

Pro Phe His Asp Ser Ala Thr Gly Pro Glu Leu Ala Ala Ser Glu Glu  
 245 250 255

Pro Phe Arg Gln Ala Val Ala Arg Ser Leu Ser His Gly Ile Asp Ser  
 260 265 270

Ser Ser Ile Ser Glu Ala Val Gly Trp Phe Glu Glu Ala Lys Ala Arg  
 275 280 285

Tyr Arg Leu Arg Ala Thr Arg Val Pro Leu Ser Arg Val Asp Lys Trp  
 290 295 300

Tyr Arg Thr Asp Thr Glu Ile Ala His Gln Asp Gly Lys Tyr Phe Ala  
 305 310 315 320

Val Ile Ala Val Ser Val Ser Ala Thr Asn Arg Glu Val Ala Ser Trp  
 325 330 335

Thr Gln Pro Met Ile Glu Pro Arg Glu Gln Gly Glu Ile Ala Leu Leu  
 340 345 350

Val Lys Arg Ile Gly Gly Val Leu His Gly Leu Val His Ala Arg Val  
 355 360 365

Glu Ala Gly Tyr Lys Trp Thr Ala Glu Ile Ala Pro Thr Val Gln Cys

370

375

380

Ser Val Ala Asn Tyr Gln Ser Thr Pro Ser Asn Asp Trp Pro Pro Phe  
 385 390 395 400

Leu Asp Asp Val Leu Thr Ala Asp Pro Glu Thr Val Arg Tyr Glu Ser  
 405 410 415

Ile Leu Ser Glu Glu Gly Gly Arg Phe Tyr Gln Ala Gln Asn Arg Tyr  
 420 425 430

Arg Ile Ile Glu Val His Glu Asp Phe Ala Ala Arg Pro Pro Ser Asp  
 435 440 445

Phe Arg Trp Met Thr Leu Gly Gln Leu Gly Glu Leu Leu Arg Ser Thr  
 450 455 460

His Phe Leu Asn Ile Gln Ala Arg Ser Leu Val Ala Ser Leu His Ser  
 465 470 475 480

Leu Trp Ala Leu Gly Arg  
 485

&lt;210&gt; 27

&lt;211&gt; 524

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(438)

&lt;223&gt; ORF11; Thioesterase

&lt;400&gt; 27

gtg agc aac gtg tgg ccg gaa aca tgg acg ccg ggg ttt ggc agg tgt 48  
 Val Ser Asn Val Trp Pro Glu Thr Trp Thr Pro Gly Phe Gly Arg Cys  
 1 5 10 15

tca tcg ctg ttg cgt cga ctc gga ttc cgc cgt gac cgg gac gat gcc 96  
 Ser Ser Leu Leu Arg Arg Leu Gly Phe Arg Arg Asp Arg Asp Asp Ala  
 20 25 30

agg cga gtc ccg aag tca gat tct tgt cca gaa tcg tcc aat ggg gtg 144  
 Arg Arg Val Pro Lys Ser Asp Ser Cys Pro Glu Ser Ser Asn Gly Val  
 35 40 45

ttg atc tcc cca gag gtt tgc gct cca acc gat ttc cga cga gga tcg 192  
 Leu Ile Ser Pro Glu Val Cys Ala Pro Thr Asp Phe Arg Arg Gly Ser  
 50 55 60

tgg cgc ccg ctg agc aac gac tac cgt gcg gtc gag aca tac cgc tgt 240  
 Trp Arg Pro Leu Ser Asn Asp Tyr Arg Ala Val Glu Thr Tyr Arg Cys  
 65 70 75 80

gcg cca gga gcg aag gtg ggt tgc ccg atc acc gtg ctg gtg gta gat 288  
 Ala Pro Gly Ala Lys Val Gly Cys Pro Ile Thr Val Leu Val Val Asp  
 85 90 95

gcc gag ccg aag gtc acc ttg gat gag gcg gaa gcc tgg cga gag cac 336  
 Ala Glu Pro Lys Val Thr Leu Asp Glu Ala Glu Ala Trp Arg Glu His  
 100 105 110

acc gag gcc gtg gcc gac gtc cgt gtc ttc tcc gcc ggg cat ttc ttc 384  
 Thr Glu Ala Val Ala Asp Val Arg Val Phe Ser Gly Gly His Phe Phe  
 115 120 125

atg acc gaa cgc cag gac gag gtg ctc gcg gtc ctt acg gcc gga tcg 432  
 Met Thr Glu Arg Gln Asp Glu Val Leu Ala Val Leu Thr Gly Gly Ser  
 130 135 140

ctt cga tgatcctcgc caggcogctg gaccagaccg cgacgccctt gggagccggc 488  
 Leu Arg  
 145

gtgcacatcg tcacggcagt gagggattgg gcatga 524

<210> 28

<211> 146

<212> PRT

<213> Saccharopolyspora spinosa

<400> 28

Val Ser Asn Val Trp Pro Glu Thr Trp Thr Pro Gly Phe Gly Arg Cys  
 1 5 10 15

Ser Ser Leu Leu Arg Arg Leu Gly Phe Arg Arg Asp Arg Asp Asp Ala  
 20 25 30

Arg Arg Val Pro Lys Ser Asp Ser Cys Pro Glu Ser Ser Asn Gly Val  
 35 40 45

Leu Ile Ser Pro Glu Val Cys Ala Pro Thr Asp Phe Arg Arg Gly Ser  
 50 55 60

Trp Arg Pro Leu Ser Asn Asp Tyr Arg Ala Val Glu Thr Tyr Arg Cys  
65 70 75 80

Ala Pro Gly Ala Lys Val Gly Cys Pro Ile Thr Val Leu Val Val Asp  
85 90 95

Ala Glu Pro Lys Val Thr Leu Asp Glu Ala Glu Ala Trp Arg Glu His  
100 105 110

Thr Glu Ala Val Ala Asp Val Arg Val Phe Ser Gly Gly His Phe Phe  
115 120 125

Met Thr Glu Arg Gln Asp Glu Val Leu Ala Val Leu Thr Gly Gly Ser  
130 135 140

Leu Arg  
45

<210> 29  
<211> 1320  
<212> DNA  
<213> Saccharopolyspora spinosa

<220>  
<221> CDS  
<222> (1)..(1317)  
<223> ORF12; Glycosyltransferase

<400> 29  
atg cgt gtc ctg ttc acc ccg ctg ccg gcg agt tcg cac ttc ttc aac 48  
Met Arg Val Leu Phe Thr Pro Leu Pro Ala Ser Ser His Phe Phe Asn  
1 5 10 15

ctg gtg ccg ttg gcg tgg gcg ttg cgt gcc gcg ggg cac gag gtc cgt 96  
Leu Val Pro Leu Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Arg  
20 25 30

gtc gcc atc tgc ccg aat atg gtg tcg atg gtc acc gga gca gga ctc 144  
Val Ala Ile Cys Pro Asn Met Val Ser Met Val Thr Gly Ala Gly Leu  
35 40 45

acc gcg gtt ccc gtc ggc gac gag ctc gac ctc atc tcc ttg gcg gcc 192  
Thr Ala Val Pro Val Gly Asp Glu Leu Asp Leu Ile Ser Leu Ala Ala  
50 55 60

aag aac gaa ctc ggc agc ggg gtc tcg ttc gag aag ggg 240  
Lys Asn Glu Leu Val Leu Gly Ser Gly Val Ser Phe Asp Glu Lys Gly  
65 70 75 80

cgg cat ccg gaa ctc ttc gac gag ctg ctg tca atc aac tcc ggc aga 288  
Arg His Pro Glu Leu Phe Asp Glu Leu Leu Ser Ile Asn Ser Gly Arg  
85 90 95

gac acg gac gcc gtg gag caa ctc cac ctt gtg gat gac cga tcg ctg 336  
Asp Thr Asp Ala Val Glu Gln Leu His Leu Val Asp Asp Arg Ser Leu  
100 105 110

gac gat ctc atg ggg ttc gcc gag aaa tgg cag cct gat ctc gtt gtg 384  
Asp Asp Leu Met Gly Phe Ala Glu Lys Trp Gln Pro Asp Leu Val Val  
115 120 125

tgg gac gct atg gtg tgt tcg ggg cca gtt gtg gcg cga gcg ctc ggc 432  
Trp Asp Ala Met Val Cys Ser Gly Pro Val Val Ala Arg Ala Leu Gly  
130 135 140

gca cga cac gtg cgg atg ctc gtc gcc ctc gat gtg tcg ggg tgg ctg 480  
Ala Arg His Val Arg Met Leu Val Ala Leu Asp Val Ser Gly Trp Leu  
145 150 155 160

cgg tcc ggt ttc ctc gaa tac cag gaa tcg aag ccg cct gag cag cgc 528  
Arg Ser Gly Phe Leu Glu Tyr Gln Glu Ser Lys Pro Pro Glu Gln Arg  
165 170 175

gtc gac ccg ctc ggg acg tgg ctg gga gcg aag ctc gcc aag ttc gga 576  
Val Asp Pro Leu Gly Thr Trp Leu Gly Ala Lys Leu Ala Lys Phe Gly  
180 185 190

gcc acg ttc gat gaa gag atc gtg acg ggc caa gcg acc ata gat ccg 624  
Ala Thr Phe Asp Glu Glu Ile Val Thr Gly Gln Ala Thr Ile Asp Pro  
195 200 205

att cca tcc tgg atg cgc ctg cct gtg gac ttg gac tac atc tcg atg 672  
Ile Pro Ser Trp Met Arg Leu Pro Val Asp Leu Asp Tyr Ile Ser Met  
210 215 220

cgt ttc gtg ccg tac aac ggt ccg gcg gtg ttg ccg gag tgg ttg cgc 720  
Arg Phe Val Pro Tyr Asn Gly Pro Ala Val Leu Pro Glu Trp Leu Arg  
225 230 235 240

gaa cga ccg acg aag ccg cgc gtc tgc atc acg cgc ggg ctg acc aag 768  
Glu Arg Pro Thr Lys Pro Arg Val Cys Ile Thr Arg Gly Leu Thr Lys  
245 250 255

cgg cgg ctg agc agg gtg acc gaa cag tac ggg gag caa agt gac cag 816  
 Arg Arg Leu Ser Arg Val Thr Glu Gln Tyr Gly Glu Gln Ser Asp Gln  
 260 265 270

gaa caa gca atg gtg gaa agg ttg ttg cgc ggc gcg gcc agg ctc gac 864  
 Glu Gln Ala Met Val Glu Arg Leu Leu Arg Gly Ala Ala Arg Leu Asp  
 275 280 285

gtc gag gtg atc gcc acc ttg tct gac gac gaa gta cgg gag atg ggg 912  
 Val Glu Val Ile Ala Thr Leu Ser Asp Asp Glu Val Arg Glu Met Gly  
 290 295 300

gag ttg ccc tcg aac gtc cgg gtc cac gaa tac gta ccg ctc aac gaa 960  
 Glu Leu Pro Ser Asn Val Arg Val His Glu Tyr Val Pro Leu Asn Glu  
 305 310 315 320

ctg ctg gag tcg tgt tca gtg atc atc cat cat ggc tcg acg acg acg -1008  
 Leu Leu Glu Ser Cys Ser Val Ile Ile His His Gly Ser Thr Thr Thr  
 325 330 335

cag gaa acc gcc acg gtc aac ggc gta ccg cag ttg att ctc cct ggg 1056  
 Gln Glu Thr Ala Thr Val Asn Gly Val Pro Gln Leu Ile Leu Pro Gly  
 340 345 350

acc ttc tgg gac gaa tct cgt agg gcg gag ctc cta gcc gat cgg gga 1104  
 Thr Phe Trp Asp Glu Ser Arg Arg Ala Glu Leu Leu Ala Asp Arg Gly  
 355 360 365

gcc ggt ctg gtc ctc gac ccc gcg acg ttt acc gaa gac gac gtg cga 1152  
 Ala Gly Leu Val Leu Asp Pro Ala Thr Phe Thr Glu Asp Asp Val Arg  
 370 375 380

ggt cag ctg gcc cgc ctg ctc gac gag ccg tcg ttc gct gcc aac gcg 1200  
 Gly Gln Leu Ala Arg Leu Leu Asp Glu Pro Ser Phe Ala Ala Asn Ala  
 385 390 395 400

gcg ctg atc cgc cgt gaa atc gag gaa agt ccc agc ccg cac gac atc 1248  
 Ala Leu Ile Arg Arg Glu Ile Glu Glu Ser Pro Ser Pro His Asp Ile  
 405 410 415

gtt cca cgt ctg gaa aag cta gtt gcc gaa cgt gag aac cgc cgc act 1296  
 Val Pro Arg Leu Glu Lys Leu Val Ala Glu Arg Glu Asn Arg Arg Thr  
 420 425 430

ggg cag tct gat ggc cat ccg tga 1320  
 Gly Gln Ser Asp Gly His Pro  
 435

<210> 30  
<211> 439  
<212> PRT  
<213> Saccharopolyspora spinosa

<400> 30  
Met Arg Val Leu Phe Thr Pro Leu Pro Ala Ser Ser His Phe Phe Asn  
1 5 10 15  
Leu Val Pro Leu Ala Trp Ala Leu Arg Ala Ala Gly His Glu Val Arg  
20 25 30  
Val Ala Ile Cys Pro Asn Met Val Ser Met Val Thr Gly Ala Gly Leu  
35 40 45  
Thr Ala Val Pro Val Gly Asp Glu Leu Asp Leu Ile Ser Leu Ala Ala  
50 55 60  
Lys Asn Glu Leu Val Leu Gly Ser Gly Val Ser Phe Asp Glu Lys Gly  
65 70 75 80  
Arg His Pro Glu Leu Phe Asp Glu Leu Leu Ser Ile Asn Ser Gly Arg  
85 90 95  
Asp Thr Asp Ala Val Glu Gln Leu His Leu Val Asp Asp Arg Ser Leu  
100 105 110  
Asp Asp Leu Met Gly Phe Ala Glu Lys Trp Gln Pro Asp Leu Val Val  
115 120 125  
Trp Asp Ala Met Val Cys Ser Gly Pro Val Val Ala Arg Ala Leu Gly  
130 135 140  
Ala Arg His Val Arg Met Leu Val Ala Leu Asp Val Ser Gly Trp Leu  
145 150 155 160  
Arg Ser Gly Phe Leu Glu Tyr Gln Glu Ser Lys Pro Pro Glu Gln Arg  
165 170 175  
Val Asp Pro Leu Gly Thr Trp Leu Gly Ala Lys Leu Ala Lys Phe Gly  
180 185 190  
Ala Thr Phe Asp Glu Glu Ile Val Thr Gly Gln Ala Thr Ile Asp Pro  
195 200 205  
Ile Pro Ser Trp Met Arg Leu Pro Val Asp Leu Asp Tyr Ile Ser Met  
210 215 220



Arg Phe Val Pro Tyr Asn Gly Pro Ala Val Leu Pro Glu Trp Leu Arg  
225 230 235 240

Glu Arg Pro Thr Lys Pro Arg Val Cys Ile Thr Arg Gly Leu Thr Lys  
245 250 255

Arg Arg Leu Ser Arg Val Thr Glu Gln Tyr Gly Glu Gln Ser Asp Gln  
260 265 270

Glu Gln Ala Met Val Glu Arg Leu Leu Arg Gly Ala Ala Arg Leu Asp  
275 280 285

Val Glu Val Ile Ala Thr Leu Ser Asp Asp Glu Val Arg Glu Met Gly  
290 295 300

Glu Leu Pro Ser Asn Val Arg Val His Glu Tyr Val Pro Leu Asn Glu  
305 310 315 320

Leu Leu Glu Ser Cys Ser Val Ile Ile His His Gly Ser Thr Thr Thr  
325 330 335

Gln Glu Thr Ala Thr Val Asn Gly Val Pro Gln Leu Ile Leu Pro Gly  
340 345 350

Thr Phe Trp Asp Glu Ser Arg Arg Ala Glu Leu Leu Ala Asp Arg Gly  
355 360 365

Ala Gly Leu Val Leu Asp Pro Ala Thr Phe Thr Glu Asp Asp Val Arg  
370 375 380

Gly Gln Leu Ala Arg Leu Leu Asp Glu Pro Ser Phe Ala Ala Asn Ala  
385 390 395 400

Ala Leu Ile Arg Arg Glu Ile Glu Glu Ser Pro Ser Pro His Asp Ile  
405 410 415

Val Pro Arg Leu Glu Lys Leu Val Ala Glu Arg Glu Asn Arg Arg Thr  
420 425 430

Gly Gln Ser Asp Gly His Pro  
435

<210> 31

<211> 1389

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(1386)

<223> ORF13; 3,4-Dehydratase

<400> 31

atg cag agc cgg aaa acc aga gcg ctg ggg aaa ggg cgc gcc aga gtg 48  
Met Gln Ser Arg Lys Thr Arg Ala Leu Gly Lys Gly Arg Ala Arg Val  
1 5 10 15

act tcg tgt gac gac act tgc gct acc gct act gag atg gtg ccg gat 96  
Thr Ser Cys Asp Asp Thr Cys Ala Thr Ala Thr Glu Met Val Pro Asp  
20 25 30

gcc aag gac cgg ata ttg gca tcc gta cgc gat tac cac cgc gaa cag -144  
Ala Lys Asp Arg Ile Leu Ala Ser Val Arg Asp Tyr His Arg Glu Gln  
35 40 45

gaa tcc ccg acc ttc gtg gct gga tcg acg ccg atc cgg cca tcg ggc 192  
Glu Ser Pro Thr Phe Val Ala Gly Ser Thr Pro Ile Arg Pro Ser Gly  
50 55 60

gcc gtg ctc gac gag gac gac cgg gtg gca ctg gtg gaa gcc gcg ctg 240  
Ala Val Leu Asp Glu Asp Asp Arg Val Ala Leu Val Glu Ala Ala Leu  
65 70 75 80

gag ctc cgg atc gcc gcg ggc ggg aat gca cgg cga ttc gag agc gag 288  
Glu Leu Arg Ile Ala Ala Gly Gly Asn Ala Arg Arg Phe Glu Ser Glu  
85 90 95

ttc gcc cgc ttc ttc ggc ctc cgc aag gct cat ctc gtc aac tcc ggt 336  
Phe Ala Arg Phe Phe Gly Leu Arg Lys Ala His Leu Val Asn Ser Gly  
100 105 110

tcg tcg gcc aat ctc ctg gca ctg agt tcg ctt acc tcc ccc aaa ctc 384  
Ser Ser Ala Asn Leu Leu Ala Leu Ser Ser Leu Thr Ser Pro Lys Leu  
115 120 125

ggc gag gca cga ctg cgg ccc ggc gac gaa gtg atc act gcg gcg gtc 432  
Gly Glu Ala Arg Leu Arg Pro Gly Asp Glu Val Ile Thr Ala Ala Val  
130 135 140

ggc ttc ccc acg acg atc aat ccg gcg gtc caa aac gga ctc gtc ccg 480  
Gly Phe Pro Thr Thr Ile Asn Pro Ala Val Gln Asn Gly Leu Val Pro  
145 150 155 160

gta ttc gtc gac gtg gaa ctg ggc acc tac aac gca acg cca gac cgc	528
Val Phe Val Asp Val Glu Leu Gly Thr Tyr Asn Ala Thr Pro Asp Arg	
165 170 175	
atc aag gcc gcc gtc acg gaa cgg acg cga gcc atc atg ctg gcg cac	576
Ile Lys Ala Ala Val Thr Glu Arg Thr Arg Ala Ile Met Leu Ala His	
180 185 190	
acc ctg ggc aac ccc ttc gcc gct gac gaa atc gcg gag atc gca aaa	624
Thr Leu Gly Asn Pro Phe Ala Ala Asp Glu Ile Ala Glu Ile Ala Lys	
195 200 205	
gaa cac gag ctg ttc ctc gtc gaa gac aac tgt gat gcg gtg gga tcc	672
Glu His Glu Leu Phe Leu Val Glu Asp Asn Cys Asp Ala Val Gly Ser	
210 215 220	
acc tac cgg gga cgg ctg acc gga acc ttc ggc gac ctg aca acg gtc	720
Thr Tyr Arg Gly Arg Leu Thr Gly Thr Phe Gly Asp Leu Thr Thr Val	
225 230 235 240	
agc ttc tat cct gcc cat cac atc acc agc ggc gag ggt ggc tgc gtg	768
Ser Phe Tyr Pro Ala His His Ile Thr Ser Gly Glu Gly Gly Cys Val	
245 250 255	
ttg acc ggc agc ctg gaa ttg gct cgc atc atc gag tcg ctg cgt gac	816
Leu Thr Gly Ser Leu Glu Leu Ala Arg Ile Ile Glu Ser Leu Arg Asp	
260 265 270	
tgg gga cgg gat tgc tgg tgc gag ccc ggc gtg gac aac acc tgc cgc	864
Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Val Asp Asn Thr Cys Arg	
275 280 285	
aag agg ttc gac tac cac ctc ggt acc ctt cca ccg ggc tac gac cac	912
Lys Arg Phe Asp Tyr His Leu Gly Thr Leu Pro Pro Gly Tyr Asp His	
290 295 300	
aag tac acg ttc tcc cac gtc ggt tac aac ctc aag acc acc gac ctg	960
Lys Tyr Thr Phe Ser His Val Gly Tyr Asn Leu Lys Thr Thr Asp Leu	
305 310 315 320	
cag gcc gca ctt gcg ctg agc cag ttg agc aag att tcc gca ttc ggg	1008
Gln Ala Ala Leu Ala Leu Ser Gln Leu Ser Lys Ile Ser Ala Phe Gly	
325 330 335	
tcg gca cgc cgc cgt aac tgg cga cgg ttg cgc gaa ggg ctg tcc ggg	1056
Ser Ala Arg Arg Arg Asn Trp Arg Arg Leu Arg Glu Gly Leu Ser Gly	
340 345 350	

ttg ccg ggc ctg ctg ctg ccg gta gcc aca ccg cac agc gac ccg agc 1104  
 Leu Pro Gly Leu Leu Leu Pro Val Ala Thr Pro His Ser Asp Pro Ser  
 355 360 365

tgg ttc ggg ttt gcg atc acc atc agt gcg gac gcc ggg ttc acc cgt 1152  
 Trp Phe Gly Phe Ala Ile Thr Ile Ser Ala Asp Ala Gly Phe Thr Arg  
 370 375 380

gcc gcc ctg gtg aac ttc ctg gaa tcc cgc aac atc ggc acc cga ctg 1200  
 Ala Ala Leu Val Asn Phe Leu Glu Ser Arg Asn Ile Gly Thr Arg Leu  
 385 390 395 400

ctg ttc ggc ggt aac atc acc cgg cac ccg gcc ttc gag cag gtg cgg 1248  
 Leu Phe Gly Gly Asn Ile Thr Arg His Pro Ala Phe Glu Gln Val Arg  
 405 410 415

tac cgg atc gcc gac gcg ctc acc aac agc gac atc gtc acc gac cga 1296  
 Tyr Arg Ile Ala Asp Ala Leu Thr Asn Ser Asp Ile Val Thr Asp Arg  
 420 425 430

acc ttc tgg gtc ggc gtc tac cca ggc ata acg gac caa atg atc gac 1344  
 Thr Phe Trp Val Gly Val Tyr Pro Gly Ile Thr Asp Gln Met Ile Asp  
 435 440 445

tac gtc gtc gaa tca atc gct gaa ttc gtg gcc aag agt tcc tag 1389  
 Tyr Val Val Glu Ser Ile Ala Glu Phe Val Ala Lys Ser Ser  
 450 455 460

<210> 32

<211> 462

<212> PRT

<213> *Saccharopolyspora spinosa*

<400> 32

Met Gln Ser Arg Lys Thr Arg Ala Leu Gly Lys Gly Arg Ala Arg Val  
 1 5 10 15

Thr Ser Cys Asp Asp Thr Cys Ala Thr Ala Thr Glu Met Val Pro Asp  
 20 25 30

Ala Lys Asp Arg Ile Leu Ala Ser Val Arg Asp Tyr His Arg Glu Gln  
 35 40 45

Glu Ser Pro Thr Phe Val Ala Gly Ser Thr Pro Ile Arg Pro Ser Gly  
 50 55 60

Ala Val Leu Asp Glu Asp Asp Arg Val Ala Leu Val Glu Ala Ala Leu

65

70

75

80

Glu Leu Arg Ile Ala Ala Gly Gly Asn Ala Arg Arg Phe Glu Ser Glu  
85 90 95

Phe Ala Arg Phe Phe Gly Leu Arg Lys Ala His Leu Val Asn Ser Gly  
100 105 110

Ser Ser Ala Asn Leu Leu Ala Leu Ser Ser Leu Thr Ser Pro Lys Leu  
115 120 125

Gly Glu Ala Arg Leu Arg Pro Gly Asp Glu Val Ile Thr Ala Ala Val  
130 135 140

Gly Phe Pro Thr Thr Ile Asn Pro Ala Val Gln Asn Gly Leu Val Pro  
145 150 155 160

Val Phe Val Asp Val Glu Leu Gly Thr Tyr Asn Ala Thr Pro Asp Arg  
165 170 175

Ile Lys Ala Ala Val Thr Glu Arg Thr Arg Ala Ile Met Leu Ala His  
180 185 190

Thr Leu Gly Asn Pro Phe Ala Ala Asp Glu Ile Ala Glu Ile Ala Lys  
195 200 205

Glu His Glu Leu Phe Leu Val Glu Asp Asn Cys Asp Ala Val Gly Ser  
210 215 220

Thr Tyr Arg Gly Arg Leu Thr Gly Thr Phe Gly Asp Leu Thr Thr Val  
225 230 235 240

Ser Phe Tyr Pro Ala His His Ile Thr Ser Gly Glu Gly Gly Cys Val  
245 250 255

Leu Thr Gly Ser Leu Glu Leu Ala Arg Ile Ile Glu Ser Leu Arg Asp  
260 265 270

Trp Gly Arg Asp Cys Trp Cys Glu Pro Gly Val Asp Asn Thr Cys Arg  
275 280 285

Lys Arg Phe Asp Tyr His Leu Gly Thr Leu Pro Pro Gly Tyr Asp His  
290 295 300

Lys Tyr Thr Phe Ser His Val Gly Tyr Asn Leu Lys Thr Thr Asp Leu  
305 310 315 320

Gln Ala Ala Leu Ala Leu Ser Gln Leu Ser Lys Ile Ser Ala Phe Gly

323

330

335

Ser Ala Arg Arg Arg Asn Trp Arg Arg Leu Arg Glu Gly Leu Ser Gly  
 340 345 350

Leu Pro Gly Leu Leu Leu Pro Val Ala Thr Pro His Ser Asp Pro Ser  
 355 360 365

Trp Phe Gly Phe Ala Ile Thr Ile Ser Ala Asp Ala Gly Phe Thr Arg  
 370 375 380

Ala Ala Leu Val Asn Phe Leu Glu Ser Arg Asn Ile Gly Thr Arg Leu  
 385 390 395 400

Leu Phe Gly Gly Asn Ile Thr Arg His Pro Ala Phe Glu Gln Val Arg  
 405 410 415

Tyr Arg Ile Ala Asp Ala Leu Thr Asn Ser Asp Ile Val Thr Asp Arg  
 420 425 430

Thr Phe Trp Val Gly Val Tyr Pro Gly Ile Thr Asp Gln Met Ile Asp  
 435 440 445

Tyr Val Val Glu Ser Ile Ala Glu Phe Val Ala Lys Ser Ser  
 450 455 460

&lt;210&gt; 33

&lt;211&gt; 1158

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1155)

&lt;223&gt; ORF14; 4-Aminotransferase

&lt;400&gt; 33

gtg atc aac ctg cac cag ccg atc ctc ggc acc gaa gaa ctc gac gcg 48  
 Val Ile Asn Leu His Gln Pro Ile Leu Gly Thr Glu Glu Leu Asp Ala  
 1 5 10 15

atc gcg gag gtg ttc gcc tcc aac tgg atc ggg ctc ggg ccg cgc acc 96  
 Ile Ala Glu Val Phe Ala Ser Asn Trp Ile Gly Leu Gly Pro Arg Thr  
 20 25 30

cgg acg ttc gag gcc gaa ttc gcc cac cac ctg gga gtg gat ccc gaa 144

Arg Thr Phe Glu Ala Glu Phe Ala His His Leu Gly Val Asp Pro Glu  
 35 40 45

cag gtc gtg ttc ctc aac tcg ggg act gcc gcg ctg ttc ctt acc gtg 192  
 Gln Val Val Phe Leu Asn Ser Gly Thr Ala Ala Leu Phe Leu Thr Val  
 50 55 60

cag gtg ctc gac ctc ggc cca ggc gac gac gtg gta ctt cct tcg ata 240  
 Gln Val Leu Asp Leu Gly Pro Gly Asp Asp Val Val Leu Pro Ser Ile  
 65 70 75 80

agc ttc gtg gcg gcg gcc aac gcc atc gca tcc tcc ggt gcc cgc ccg 288  
 Ser Phe Val Ala Ala Ala Asn Ala Ile Ala Ser Ser Gly Ala Arg Pro  
 85 90 95

gtg ttc tgc gac gtc gac ccc cgg acg ttg aac ccc acg ctg gat gat 336  
 Val Phe Cys Asp Val Asp Pro Arg Thr Leu Asn Pro Thr Leu Asp Asp  
 100 105 110

gtg gcg agg gcc atc acg ccg gcg acc aag gcc gta ttg ctg ctc cac 384  
 Val Ala Arg Ala Ile Thr Pro Ala Thr Lys Ala Val Leu Leu Leu His  
 115 120 125

tat gga gga tcg ccg gga gaa gtc acc gcg atc gcc gat ttc tgc cgt 432  
 Tyr Gly Gly Ser Pro Gly Glu Val Thr Ala Ile Ala Asp Phe Cys Arg  
 130 135 140

gaa aag ggc ctc atg ctc atc gag gac tcc gcc tgc gcg gtg gca tcg 480  
 Glu Lys Gly Leu Met Leu Ile Glu Asp Ser Ala Cys Ala Val Ala Ser  
 145 150 155 160

tcc gtg cac ggc acc gct tgc gga acc ttt ggt gac ctg gcc acg tgg 528  
 Ser Val His Gly Thr Ala Cys Gly Thr Phe Gly Asp Leu Ala Thr Trp  
 165 170 175

agt ttc gat gcg atg aag atc ctg gtc acc ggg gat ggg ggc atg ttc 576  
 Ser Phe Asp Ala Met Lys Ile Leu Val Thr Gly Asp Gly Gly Met Phe  
 180 185 190

tac gcg gcg gat ccg gag ctg gcg cac cgc gca aga cga ctc gcc tac 624  
 Tyr Ala Ala Asp Pro Glu Leu Ala His Arg Ala Arg Arg Leu Ala Tyr  
 195 200 205

cac ggt ctt gag cag atg agc gga ttc gat tcg gcc aag tct tcc aac 672  
 His Gly Leu Glu Gln Met Ser Gly Phe Asp Ser Ala Lys Ser Ser Asn  
 210 215 220

cgc tgg tgg gat att cgc gtc gaa gac atc ggc cag cgg ctg atc ggg 720





<400> 34

Val Ile Asn Leu His Gln Pro Ile Leu Gly Thr Glu Glu Leu Asp Ala  
1 5 10 15

Ile Ala Glu Val Phe Ala Ser Asn Trp Ile Gly Leu Gly Pro Arg Thr  
20 25 30

Arg Thr Phe Glu Ala Glu Phe Ala His His Leu Gly Val Asp Pro Glu  
35 40 45

Gln Val Val Phe Leu Asn Ser Gly Thr Ala Ala Leu Phe Leu Thr Val  
50 55 60

Gln Val Leu Asp Leu Gly Pro Gly Asp Asp Val Val Leu Pro Ser Ile  
65 70 75 80

Ser Phe Val Ala Ala Ala Asn Ala Ile Ala Ser Ser Gly Ala Arg Pro  
85 90 95

Val Phe Cys Asp Val Asp Pro Arg Thr Leu Asn Pro Thr Leu Asp Asp  
100 105 110

Val Ala Arg Ala Ile Thr Pro Ala Thr Lys Ala Val Leu Leu Leu His  
115 120 125

Tyr Gly Gly Ser Pro Gly Glu Val Thr Ala Ile Ala Asp Phe Cys Arg  
130 135 140

Glu Lys Gly Leu Met Leu Ile Glu Asp Ser Ala Cys Ala Val Ala Ser  
145 150 155 160

Ser Val His Gly Thr Ala Cys Gly Thr Phe Gly Asp Leu Ala Thr Trp  
165 170 175

Ser Phe Asp Ala Met Lys Ile Leu Val Thr Gly Asp Gly Gly Met Phe  
180 185 190

Tyr Ala Ala Asp Pro Glu Leu Ala His Arg Ala Arg Arg Leu Ala Tyr  
195 200 205

His Gly Leu Glu Gln Met Ser Gly Phe Asp Ser Ala Lys Ser Ser Asn  
210 215 220

Arg Trp Trp Asp Ile Arg Val Glu Asp Ile Gly Gln Arg Leu Ile Gly  
225 230 235 240

Asn Asp Met Thr Ala Ala Leu Gly Ser Val Gln Leu Arg Lys Leu Pro

245

250

255

Glu Phe Ile Asn Arg Arg Arg Glu Ile Ala Thr Gln Tyr Asp Arg Leu  
 260 265 270

Leu Ser Asp Val Pro Gly Val Leu Leu Pro Pro Thr Leu Pro Asp Gly  
 275 280 285

His Val Ser Ser His Tyr Phe Tyr Trp Val Gln Leu Ala Pro Glu Ile  
 290 295 300

Arg Asp Gln Val Ala Gln Gln Met Leu Glu Arg Gly Ile Tyr Thr Ser  
 305 310 315 320

Tyr Arg Tyr Pro Pro Leu His Lys Val Pro Ile Tyr Arg Ala Asp Cys  
 325 330 335

Lys Leu Pro Ser Ala Glu His Ala Cys Arg Arg Thr Leu Leu Leu Pro  
 340 345 350

Leu His Pro Ser Leu Asp Asp Ala Glu Val Arg Thr Val Ala Asp Glu  
 355 360 365

Phe Gln Lys Ala Val Glu His His Ile Ser Gln Arg Ser Pro Leu Arg  
 370 375 380

Lys  
 385

&lt;210&gt; 35

&lt;211&gt; 750

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(747)

&lt;223&gt; ORF15; N-Dimethyltransferase

&lt;400&gt; 35

atg tcg cgc gtg agc gac aca ttc gca gaa acc tcc tcg gta tac agc 48  
 Met Ser Arg Val Ser Asp Thr Phe Ala Glu Thr Ser Ser Val Tyr Ser  
 1 5 10 15

cca gat cat gcc gac atc tac gac gcg atc cac tcc gcg cgt ggc cgg 96  
 Pro Asp His Ala Asp Ile Tyr Asp Ala Ile His Ser Ala Arg Gly Arg

20	25	30	
gac tgg gca gcc gag gcc ggg gaa gta gtc cag ctc gta cgc acc agg			144
Asp Trp Ala Ala Glu Ala Gly Glu Val Val Gln Leu Val Arg Thr Arg			
35	40	45	
ctg ccc gaa gca cag tcc cta ctc gac gtc gcc tgt ggg acc ggg gcg			192
Leu Pro Glu Ala Gln Ser Leu Leu Asp Val Ala Cys Gly Thr Gly Ala			
50	55	60	
cac cta gag cga ttc cgt gcc gaa tac gcg aag gtc gcg ggg ctt gaa			240
His Leu Glu Arg Phe Arg Ala Glu Tyr Ala Lys Val Ala Gly Leu Glu			
65	70	75	80
ctg tcc gat gcg atg cgg gag atc gcg atc aga cga gtc cct gag gta			288
Leu Ser Asp Ala Met Arg Glu Ile Ala Ile Arg Arg Val Pro Glu Val			
85	90	95	
ccg att cac atc ggt gac atc cgc gat ttc gac ctc ggc gag cca ttc			336
Pro Ile His Ile Gly Asp Ile Arg Asp Phe Asp Leu Gly Glu Pro Phe			
100	105	110	
gac gtc atc acc tgc ctg tgc ttt acc gcg gct tac atg cgg acc gtt			384
Asp Val Ile Thr Cys Leu Cys Phe Thr Ala Ala Tyr Met Arg Thr Val			
115	120	125	
gac gac ctg cga cgc gtg acg cgg aac atg gcc cgg cac ctg gcc cct			432
Asp Asp Leu Arg Arg Val Thr Arg Asn Met Ala Arg His Leu Ala Pro			
130	135	140	
ggt gga gtc gcg gtc atc gaa ccc tgg tgg ttt ccc gac aag ttc atc			480
Gly Gly Val Ala Val Ile Glu Pro Trp Trp Phe Pro Asp Lys Phe Ile			
145	150	155	160
gac ggg ttc gtc acc gga gcc gtc gcg cac cac ggc gag cgg gtg atc			528
Asp Gly Phe Val Thr Gly Ala Val Ala His His Gly Glu Arg Val Ile			
165	170	175	
agc cgg cta tcg cac tcg gtc ctg gag ggc cgt acg agc cgg atg acc			576
Ser Arg Leu Ser His Ser Val Leu Glu Gly Arg Thr Ser Arg Met Thr			
180	185	190	
gtc cgc tac aca gtc gcc gaa ccc acc ggg atc cgg gac ttc aca gag			624
Val Arg Tyr Thr Val Ala Glu Pro Thr Gly Ile Arg Asp Phe Thr Glu			
195	200	205	
ttc gaa atc ctc tcg ctg ttc act gag gac gag tac acc gcc gcg ctc			672
Phe Glu Ile Leu Ser Leu Phe Thr Glu Asp Glu Tyr Thr Ala Ala Leu			

210

215

220

gaa gac gca ggg atc cgc gcg gaa tac ctt cct gga gca ccg aac ggc 720  
 Glu Asp Ala Gly Ile Arg Ala Glu Tyr Leu Pro Gly Ala Pro Asn Gly  
 225 230 235 240

cga ggc ctg ttc gtc gga atc cgc aac tga 750  
 Arg Gly Leu Phe Val Gly Ile Arg Asn  
 245

&lt;210&gt; 36

&lt;211&gt; 249

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 36

Met Ser Arg Val Ser Asp Thr Phe Ala Glu Thr Ser Ser Val Tyr Ser  
 1 5 10 15

Pro Asp His Ala Asp Ile Tyr Asp Ala Ile His Ser Ala Arg Gly Arg  
 20 25 30

Asp Trp Ala Ala Glu Ala Gly Glu Val Val Gln Leu Val Arg Thr Arg  
 35 40 45

Leu Pro Glu Ala Gln Ser Leu Leu Asp Val Ala Cys Gly Thr Gly Ala  
 50 55 60

His Leu Glu Arg Phe Arg Ala Glu Tyr Ala Lys Val Ala Gly Leu Glu  
 65 70 75 80

Leu Ser Asp Ala Met Arg Glu Ile Ala Ile Arg Arg Val Pro Glu Val  
 85 90 95

Pro Ile His Ile Gly Asp Ile Arg Asp Phe Asp Leu Gly Glu Pro Phe  
 100 105 110

Asp Val Ile Thr Cys Leu Cys Phe Thr Ala Ala Tyr Met Arg Thr Val  
 115 120 125

Asp Asp Leu Arg Arg Val Thr Arg Asn Met Ala Arg His Leu Ala Pro  
 130 135 140

Gly Gly Val Ala Val Ile Glu Pro Trp Trp Phe Pro Asp Lys Phe Ile  
 145 150 155 160

Asp Gly Phe Val Thr Gly Ala Val Ala His His Gly Glu Arg Val Ile

165                      170                      175  
 Ser Arg Leu Ser His Ser Val Leu Glu Gly Arg Thr Ser Arg Met Thr  
                          180                      185                      190  
 Val Arg Tyr Thr Val Ala Glu Pro Thr Gly Ile Arg Asp Phe Thr Glu  
                          195                      200                      205  
 Phe Glu Ile Leu Ser Leu Phe Thr Glu Asp Glu Tyr Thr Ala Ala Leu  
                          210                      215                      220  
 Glu Asp Ala Gly Ile Arg Ala Glu Tyr Leu Pro Gly Ala Pro Asn Gly  
                          225                      230                      235                      240  
 Arg Gly Leu Phe Val Gly Ile Arg Asn  
    245

<210> 37  
 <211> 726  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(723)  
 <223> ORF16; 3,4-Reduktase

<400> 37  
 atg agc gaa cag acg att gca ctg gtc acc ggc gca aac aag gga atc    48  
 Met Ser Glu Gln Thr Ile Ala Leu Val Thr Gly Ala Asn Lys Gly Ile  
       1                      5                      10                      15  
 gga tac gag atc gcg gcc ggg ctc ggc gcg ctg ggg tgg agc gtc gga    96  
 Gly Tyr Glu Ile Ala Ala Gly Leu Gly Ala Leu Gly Trp Ser Val Gly  
                          20                      25                      30  
 atc ggg gca cgg gac cac cag cgc ggg gag gat gcc gtg gcg aaa ttg    144  
 Ile Gly Ala Arg Asp His Gln Arg Gly Glu Asp Ala Val Ala Lys Leu  
                          35                      40                      45  
 cgt gcg gac ggc gtc gat gcg ttc gcg gta tcc ctg gac gtg aca gac    192  
 Arg Ala Asp Gly Val Asp Ala Phe Ala Val Ser Leu Asp Val Thr Asp  
                          50                      55                      60  
 gac gcg agc gtc gcg gct gct gcg gct ctg ctc gag gag cgc gcc ggc    240  
 Asp Ala Ser Val Ala Ala Ala Ala Ala Leu Leu Glu Glu Arg Ala Gly

65	70	75	80	
cgg ctc gat gtg ctg gtt aat aac gcc ggc atc gcc ggg gca tgg ccg				288
Arg Leu Asp Val Leu Val Asn Asn Ala Gly Ile Ala Gly Ala Trp Pro				
	85	90	95	
gag gag ccc tcg acc gtc aca ccg gcg agc ctc cgg gcg gtg gtg gag				336
Glu Glu Pro Ser Thr Val Thr Pro Ala Ser Leu Arg Ala Val Val Glu				
	100	105	110	
acc aac gtg atc ggc gtc gtt cgg gtt acc aac gct atg ctg ccg ttg				384
Thr Asn Val Ile Gly Val Val Arg Val Thr Asn Ala Met Leu Pro Leu				
	115	120	125	
cta cgc cgc tcc gag cgc ccg cgg atc gtc aac cag tcc agc cac gtc				432
Leu Arg Arg Ser Glu Arg Pro Arg Ile Val Asn Gln Ser Ser His Val				
	130	135	140	
gct tcc ctg acc ttg caa acc acg ccg ggc gtc gac ctc ggc ggg atc				480
Ala Ser Leu Thr Leu Gln Thr Thr Pro Gly Val Asp Leu Gly Gly Ile				
	145	150	155	160
agc gga gcc tac tca ccg tcg aag acg ttc ctc aac gcg atc acc atc				528
Ser Gly Ala Tyr Ser Pro Ser Lys Thr Phe Leu Asn Ala Ile Thr Ile				
	165	170	175	
cag tac gcc aag gaa ctc agc gat acc aac atc aaa atc aac aac gcc				576
Gln Tyr Ala Lys Glu Leu Ser Asp Thr Asn Ile Lys Ile Asn Asn Ala				
	180	185	190	
tgc ccc ggc tac gtc gcg acc gac ctt aac ggc ttc cac gga acc agc				624
Cys Pro Gly Tyr Val Ala Thr Asp Leu Asn Gly Phe His Gly Thr Ser				
	195	200	205	
acg ccg gca gac ggt gcc agg atc gcc att cgg ctc gcc acg ctg cca				672
Thr Pro Ala Asp Gly Ala Arg Ile Ala Ile Arg Leu Ala Thr Leu Pro				
	210	215	220	
gac gac ggc ccg acc gga ggc atg ttc gac gac gcc ggg aat gtg ccc				720
Asp Asp Gly Pro Thr Gly Gly Met Phe Asp Asp Ala Gly Asn Val Pro				
	225	230	235	240
tgg tga				726
Trp				

<210> 38

<211> 241

<212> PRT

<213> Saccharopolyspora spinosa

<400> 38

Met Ser Glu Gln Thr Ile Ala Leu Val Thr Gly Ala Asn Lys Gly Ile  
1 5 10 15

Gly Tyr Glu Ile Ala Ala Gly Leu Gly Ala Leu Gly Trp Ser Val Gly  
20 25 30

Ile Gly Ala Arg Asp His Gln Arg Gly Glu Asp Ala Val Ala Lys Leu  
35 40 45

Arg Ala Asp Gly Val Asp Ala Phe Ala Val Ser Leu Asp Val Thr Asp  
50 55 60

Asp Ala Ser Val Ala Ala Ala Ala Ala Leu Leu Glu Glu Arg Ala Gly  
65 70 75 80

Arg Leu Asp Val Leu Val Asn Asn Ala Gly Ile Ala Gly Ala Trp Pro  
85 90 95

Glu Glu Pro Ser Thr Val Thr Pro Ala Ser Leu Arg Ala Val Val Glu  
100 105 110

Thr Asn Val Ile Gly Val Val Arg Val Thr Asn Ala Met Leu Pro Leu  
115 120 125

Leu Arg Arg Ser Glu Arg Pro Arg Ile Val Asn Gln Ser Ser His Val  
130 135 140

Ala Ser Leu Thr Leu Gln Thr Thr Pro Gly Val Asp Leu Gly Gly Ile  
145 150 155 160

Ser Gly Ala Tyr Ser Pro Ser Lys Thr Phe Leu Asn Ala Ile Thr Ile  
165 170 175

Gln Tyr Ala Lys Glu Leu Ser Asp Thr Asn Ile Lys Ile Asn Asn Ala  
180 185 190

Cys Pro Gly Tyr Val Ala Thr Asp Leu Asn Gly Phe His Gly Thr Ser  
195 200 205

Thr Pro Ala Asp Gly Ala Arg Ile Ala Ile Arg Leu Ala Thr Leu Pro  
210 215 220

Asp Asp Gly Pro Thr Gly Gly Met Phe Asp Asp Ala Gly Asn Val Pro  
225 230 235 240

Trp

<210> 39  
<211> 837  
<212> DNA  
<213> Saccharopolyspora spinosa

<220>  
<221> CDS  
<222> (1)..(834)  
<223> ORF17; Transkriptions-Regulator

<400> 39

atg gag acg cgg gag ttg cgg tac ttc gtt gca gtc gcc gag gag ttg	48
Met Glu Thr Arg Glu Leu Arg Tyr Phe Val Ala Val Ala Glu Glu Leu	
1 5 10 15	
cac ttc ggc cgg gcc gcc cag cgc ctg ggc atc gcc cag ccg ccg ctg	96
His Phe Gly Arg Ala Ala Gln Arg Leu Gly Ile Ala Gln Pro Pro Leu	
20 25 30	
tcg cgg acg atc gcc cag ctc gag caa cga ctc gga gtc gtg ttg ctg	144
Ser Arg Thr Ile Ala Gln Leu Glu Gln Arg Leu Gly Val Val Leu Leu	
35 40 45	
caa cgc acc agc cgc aaa gtc tcg ctc acc gaa gcc ggg gca atg ctg	192
Gln Arg Thr Ser Arg Lys Val Ser Leu Thr Glu Ala Gly Ala Met Leu	
50 55 60	
ctg acc gaa ggc cgg gcg atc ctc ggc gcg ctg gca gca gcc gag cga	240
Leu Thr Glu Gly Arg Ala Ile Leu Gly Ala Leu Ala Ala Ala Glu Arg	
65 70 75 80	
cgc acc cag cgt gcc gcg acg agc cag ccc tcg cta gtc ctg gct gcc	288
Arg Thr Gln Arg Ala Ala Thr Ser Gln Pro Ser Leu Val Leu Ala Ala	
85 90 95	
aag gcc ggc gcc tcc ggt gag ctg ctg gcg aag ttg ctc gac gcg tac	336
Lys Ala Gly Ala Ser Gly Glu Leu Leu Ala Lys Leu Leu Asp Ala Tyr	
100 105 110	
gcc gcc gag ccg gga gcc gtg gcc gtc gac ctg ctg ctc tgc gaa tcc	384
Ala Ala Glu Pro Gly Ala Val Ala Val Asp Leu Leu Leu Cys Glu Ser	
115 120 125	



cag ccc cag aaa acg ctg cat gac ggc cgg gcc gac gtg gcg ctg ttg 432  
 Gln Pro Gln Lys Thr Leu His Asp Gly Arg Ala Asp Val Ala Leu Leu  
 130 135 140

cat caa ccc ttc gac ccg acg gcc gaa ctc gac atc gaa att ctg aac 480  
 His Gln Pro Phe Asp Pro Thr Ala Glu Leu Asp Ile Glu Ile Leu Asn  
 145 150 155 160

acc gag caa caa gtc gcc att ctt ccg acc tcg cat ccg ctt gcc agc 528  
 Thr Glu Gln Gln Val Ala Ile Leu Pro Thr Ser His Pro Leu Ala Ser  
 165 170 175

gag ccc cat gta cgg atg gcg gat gtc agc tca ctg ccg gat ctc ccg 576  
 Glu Pro His Val Arg Met Ala Asp Val Ser Ser Leu Pro Asp Leu Pro  
 180 185 190

ctt gcg cgc tgg ccc ggc ccc gac ggc gtc tat cca gat ggc ccc ggc 624  
 Leu Ala Arg Trp Pro Gly Pro Asp Gly Val Tyr Pro Asp Gly Pro Gly  
 195 200 205

gtg gaa gta cgc aac cag acg caa ctg ttc caa atg atc gca ctc ggc 672  
 Val Glu Val Arg Asn Gln Thr Gln Leu Phe Gln Met Ile Ala Leu Gly  
 210 215 220

cgc act acc gtg gtc atg ccc gaa tcc agt cgc gtc aac ctg ctc gaa 720  
 Arg Thr Thr Val Val Met Pro Glu Ser Ser Arg Val Asn Leu Leu Glu  
 225 230 235 240

ggc ctc gcc gcc gta ccg gtt cta gac gcg ccg gac gtg acg aca gtc 768  
 Gly Leu Ala Ala Val Pro Val Leu Asp Ala Pro Asp Val Thr Thr Val  
 245 250 255

atc gcc tgg ccg ccc cac agc cgc tcc cga gca ctc gcc ggc ttg gtc 816  
 Ile Ala Trp Pro Pro His Ser Arg Ser Arg Ala Leu Ala Gly Leu Val  
 260 265 270

cgc gtg gcc aca ctc ctc taa 837  
 Arg Val Ala Thr Leu Leu  
 275

<210> 40

<211> 278

<212> PRT

<213> Saccharopolyspora spinosa

<400> 40

Met Glu Thr Arg Gln Leu Arg Tyr Phe Val Ala Val Ala Glu Glu Leu  
 1 5 10 15  
 His Phe Gly Arg Ala Ala Gln Arg Leu Gly Ile Ala Gln Pro Pro Leu  
 20 25 30  
 Ser Arg Thr Ile Ala Gln Leu Glu Gln Arg Leu Gly Val Val Leu Leu  
 35 40 45  
 Gln Arg Thr Ser Arg Lys Val Ser Leu Thr Glu Ala Gly Ala Met Leu  
 50 55 60  
 Leu Thr Glu Gly Arg Ala Ile Leu Gly Ala Leu Ala Ala Ala Glu Arg  
 65 70 75 80  
 Arg Thr Gln Arg Ala Ala Thr Ser Gln Pro Ser Leu Val Leu Ala Ala  
 85 90 95  
 Lys Ala Gly Ala Ser Gly Glu Leu Leu Ala Lys Leu Leu Asp Ala Tyr  
 100 105 110  
 Ala Ala Glu Pro Gly Ala Val Ala Val Asp Leu Leu Leu Cys Glu Ser  
 115 120 125  
 Gln Pro Gln Lys Thr Leu His Asp Gly Arg Ala Asp Val Ala Leu Leu  
 130 135 140  
 His Gln Pro Phe Asp Pro Thr Ala Glu Leu Asp Ile Glu Ile Leu Asn  
 145 150 155 160  
 Thr Glu Gln Gln Val Ala Ile Leu Pro Thr Ser His Pro Leu Ala Ser  
 165 170 175  
 Glu Pro His Val Arg Met Ala Asp Val Ser Ser Leu Pro Asp Leu Pro  
 180 185 190  
 Leu Ala Arg Trp Pro Gly Pro Asp Gly Val Tyr Pro Asp Gly Pro Gly  
 195 200 205  
 Val Glu Val Arg Asn Gln Thr Gln Leu Phe Gln Met Ile Ala Leu Gly  
 210 215 220  
 Arg Thr Thr Val Val Met Pro Glu Ser Ser Arg Val Asn Leu Leu Glu  
 225 230 235 240  
 Gly Leu Ala Ala Val Pro Val Leu Asp Ala Pro Asp Val Thr Thr Val  
 245 250 255

Ile Ala Trp Pro Pro His Ser Arg Ser Arg Ala Leu Ala Gly Leu Val  
 260 265 270

Arg Val Ala Thr Leu Leu  
 275

<210> 41  
 <211> 7788  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(7785)  
 <223> ORF18; Polyketidsynthase

<400> 41  
 atg agc gaa gcc ggg aac ctg ata gcc gtc atc gga ctg tcc tgc cgc 48  
 Met Ser Glu Ala Gly Asn Leu Ile Ala Val Ile Gly Leu Ser Cys Arg  
 1 5 10 15  
 cta ccc cag gcg cct gac ccg gct tcc ttc tgg cgg ttg ctg cgc acc 96  
 Leu Pro Gln Ala Pro Asp Pro Ala Ser Phe Trp Arg Leu Leu Arg Thr  
 20 25 30  
 gga acg gac gcc atc acc acg gtc ccg gaa ggg cgg tgg ggc gac ccg 144  
 Gly Thr Asp Ala Ile Thr Thr Val Pro Glu Gly Arg Trp Gly Asp Pro  
 35 40 45  
 ttg cct ggt cgg gat gcg ccc aag ggc ccg gaa tgg ggt ggt ttc ctg 192  
 Leu Pro Gly Arg Asp Ala Pro Lys Gly Pro Glu Trp Gly Gly Phe Leu  
 50 55 60  
 gct gat gtc gac tgc ttc gat ccc gag ttc ttc ggg atc tcg ccg cga 240  
 Ala Asp Val Asp Cys Phe Asp Pro Glu Phe Phe Gly Ile Ser Pro Arg  
 65 70 75 80  
 gaa gcg gca gcc gtg gat ccc cag cag agg ctg gct ctg gag ctc gcc 288  
 Glu Ala Ala Ala Val Asp Pro Gln Gln Arg Leu Ala Leu Glu Leu Ala  
 85 90 95  
 tgg gag gca ctc gaa gac gcc ggt atc ccc gcc ggc gag ctg cgc ggt 336  
 Trp Glu Ala Leu Glu Asp Ala Gly Ile Pro Ala Gly Glu Leu Arg Gly  
 100 105 110  
 act gcc gcc ggt gtg ttc atg ggg gcg atc tct gac gac tac gcc gcc 384

Thr Ala Ala Gly Val Phe Met Gly Ala Ile Ser Asp Asn Tyr Ala Ala	
115 120 125	
ctg ctg cgc gag agc ccg ccg gaa gtg gct gcg cag tac cgc ctc acc	432
Leu Leu Arg Glu Ser Pro Pro Glu Val Ala Ala Gln Tyr Arg Leu Thr	
130 135 140	
ggc acc cat cga agc ctg atc gcc aac cgc gtg tcc tat gtg ctc ggc	480
Gly Thr His Arg Ser Leu Ile Ala Asn Arg Val Ser Tyr Val Leu Gly	
145 150 155 160	
ctg cgc ggg cca agc ctg acg gtg gat tca ggt cag tcc tcg tcc ctg	528
Leu Arg Gly Pro Ser Leu Thr Val Asp Ser Gly Gln Ser Ser Ser Leu	
165 170 175	
gtc ggc gtg cat ctc gcc agc gag agc ctg cga cgg ggt gag tgc acg	576
Val Gly Val His Leu Ala Ser Glu Ser Leu Arg Arg Gly Glu Cys Thr	
180 185 190	
atc gca ctc gcc ggc ggc gtg aac ctc aac ctg gcc gcc gag agc aac	624
Ile Ala Leu Ala Gly Gly Val Asn Leu Asn Leu Ala Ala Glu Ser Asn	
195 200 205	
agc gct ctg atg gac ttc ggc gcg ctc tcc ccg gac ggt cgc tgc ttc	672
Ser Ala Leu Met Asp Phe Gly Ala Leu Ser Pro Asp Gly Arg Cys Phe	
210 215 220	
acc ttc gat gtg cgg gcg aac ggt tac gtc cgt ggt gag ggc ggc ggc	720
Thr Phe Asp Val Arg Ala Asn Gly Tyr Val Arg Gly Glu Gly Gly Gly	
225 230 235 240	
ctt gtc gtg ctg aag aag gcc gat cag gcg cac gcc gat ggc gac cgg	768
Leu Val Val Leu Lys Lys Ala Asp Gln Ala His Ala Asp Gly Asp Arg	
245 250 255	
atc tac tgc ctc atc cgc ggc agc gcg gtc aac aac gat ggg ggc ggt	816
Ile Tyr Cys Leu Ile Arg Gly Ser Ala Val Asn Asn Asp Gly Gly Gly	
260 265 270	
gcc ggg ctc acc gtt ccg gcg gcg gac gcc cag gcg gag ctg ctg cgc	864
Ala Gly Leu Thr Val Pro Ala Ala Asp Ala Gln Ala Glu Leu Leu Arg	
275 280 285	
cag gca tac cgg aac gcg ggc gtc gac ccg gcc gcc gtg cag tat gtc	912
Gln Ala Tyr Arg Asn Ala Gly Val Asp Pro Ala Ala Val Gln Tyr Val	
290 295 300	
gag ctc cac ggc agc gcg acc agg gtc ggg gat ccc gtc gaa gca gca	960



Pro Lys Leu Leu Ala Gly Leu Lys Ala Leu Ala Asp Gly Asp Asp Ala	
500	510
ccc acg ctg tgc cac ggc gcg act tcc ggc gag cgg gca gcg gtc ttc	1584
Pro Thr Leu Cys His Gly Ala Thr Ser Gly Glu Arg Ala Ala Val Phe	
515	525
gtc ttt ccc gga cag ggc agc cag tgg atc ggg atg ggt agg cag ctg	1632
Val Phe Pro Gly Gln Gly Ser Gln Trp Ile Gly Met Gly Arg Gln Leu	
530	540
ctc gaa acc tcc gag gtt ttc gcg gcg tcg atg tcg gac tgc gcc gac	1680
Leu Glu Thr Ser Glu Val Phe Ala Ala Ser Met Ser Asp Cys Ala Asp	
545	560
gca ttg gcg ccg cac ctg gat tgg tcc ctg ctg gat gtg ctg cgc aac	1728
Ala Leu Ala Pro His Leu Asp Trp Ser Leu Leu Asp Val Leu Arg Asn	
565	575
gcg gcc ggc gct gcg cac ctt gac cac gac gat gtc gtc cag ccc gcg	1776
Ala Ala Gly Ala Ala His Leu Asp His Asp Asp Val Val Gln Pro Ala	
580	590
ctg ttc gcc atc atg gtc tcg ctc gcg gag ctc tgg cgt tcg tgg ggc	1824
Leu Phe Ala Ile Met Val Ser Leu Ala Glu Leu Trp Arg Ser Trp Gly	
595	605
gtg cgt ccg gtg gcg gtc gtc ggg cac tcg cag ggg gag atc gcg gcg	1872
Val Arg Pro Val Ala Val Val Gly His Ser Gln Gly Glu Ile Ala Ala	
610	620
gcc tgc gtc gcc ggg gcc ctg tcc gtc cgc gat gcc gcc agg gtg gtg	1920
Ala Cys Val Ala Gly Ala Leu Ser Val Arg Asp Ala Ala Arg Val Val	
625	640
gcg gtg cgc agc agg ctt ctg acg gcg ctg gcc ggc agt ggc gcg atg	1968
Ala Val Arg Ser Arg Leu Leu Thr Ala Leu Ala Gly Ser Gly Ala Met	
645	655
gcc tcg ttg cag cat ccc gcc gaa gag gtg cgg caa atc ctg ttg ccc	2016
Ala Ser Leu Gln His Pro Ala Glu Glu Val Arg Gln Ile Leu Leu Pro	
660	670
tgg cgc gat cgg atc ggc gtg gcg ggg gtg aac gga ccg tcg tcg acc	2064
Trp Arg Asp Arg Ile Gly Val Ala Gly Val Asn Gly Pro Ser Ser Thr	
675	685
ctg gtg tca ggg gac cgg gag gcg atg gcg gaa ctg ctg gcc gag tgc	2112

Leu	Val	Ser	Gly	Asp	Arg	Glu	Ala	Met	Ala	Glu	Leu	Leu	Ala	Glu	Cys	
690						695					700					
gca	gac	cga	gag	ctc	cgg	atg	cgc	cgg	att	ccc	gtt	gaa	tac	gcc	tcc	2160
Ala	Asp	Arg	Glu	Leu	Arg	Met	Arg	Arg	Ile	Pro	Val	Glu	Tyr	Ala	Ser	
705					710				715						720	
cat	tcg	cct	cac	atc	gag	gtt	gtc	cgg	gat	gag	ctg	ctg	ggg	ctg	ttg	2208
His	Ser	Pro	His	Ile	Glu	Val	Val	Arg	Asp	Glu	Leu	Leu	Gly	Leu	Leu	
				725				730						735		
gcg	ccg	gtc	gaa	ccc	agg	acg	gga	agc	atc	ccg	atc	tat	tcg	acg	acg	2256
Ala	Pro	Val	Glu	Pro	Arg	Thr	Gly	Ser	Ile	Pro	Ile	Tyr	Ser	Thr	Thr	
			740					745						750		
acc	ggg	gac	ctg	ctg	gac	cgg	ccg	atg	gac	gcc	gac	tac	tgg	tac	cgc	2304
Thr	Gly	Asp	Leu	Leu	Asp	Arg	Pro	Met	Asp	Ala	Asp	Tyr	Trp	Tyr	Arg	
		755					760					765				
aac	ctt	cgt	caa	ccg	gtg	ctg	ttc	gaa	gcg	gcc	gtc	gag	gcc	ctg	ttg	2352
Asn	Leu	Arg	Gln	Pro	Val	Leu	Phe	Glu	Ala	Ala	Val	Glu	Ala	Leu	Leu	
	770					775					780					
aag	cgg	ggg	tac	gac	gca	ttc	atc	gag	atc	agc	cca	cac	ccg	gtg	ctg	2400
Lys	Arg	Gly	Tyr	Asp	Ala	Phe	Ile	Glu	Ile	Ser	Pro	His	Pro	Val	Leu	
785					790					795					800	
act	gcg	aac	atc	cag	gaa	acc	gcc	gtg	cga	gca	ggg	cgg	gag	gta	gtg	2448
Thr	Ala	Asn	Ile	Gln	Glu	Thr	Ala	Val	Arg	Ala	Gly	Arg	Glu	Val	Val	
				805				810						815		
gcg	ctc	ggg	aca	ctc	cgc	cgc	ggc	gaa	ggg	ggc	atg	cgg	cag	gcg	ctg	2496
Ala	Leu	Gly	Thr	Leu	Arg	Arg	Gly	Glu	Gly	Gly	Met	Arg	Gln	Ala	Leu	
			820				825						830			
acg	tcg	ctg	gcc	aga	gca	cac	gta	cac	gga	gtg	gcc	gcg	gac	tgg	cac	2544
Thr	Ser	Leu	Ala	Arg	Ala	His	Val	His	Gly	Val	Ala	Ala	Asp	Trp	His	
			835				840						845			
gcg	gtc	ttc	gcc	ggg	acc	ggg	gcg	cag	cgg	gtc	gac	ctg	ccg	acg	tac	2592
Ala	Val	Phe	Ala	Gly	Thr	Gly	Ala	Gln	Arg	Val	Asp	Leu	Pro	Thr	Tyr	
	850					855				860						
gcc	ttt	cag	cga	cag	cgc	tac	tgg	ctg	gac	gcg	aag	ctt	ccc	gac	gtc	2640
Ala	Phe	Gln	Arg	Gln	Arg	Tyr	Trp	Leu	Asp	Ala	Lys	Leu	Pro	Asp	Val	
865					870				875					880		
gcc	atg	ccc	gag	agc	gac	gtg	tcg	acg	gcg	ttg	cgg	gaa	aag	ctg	cgg	2688

Ala Met Pro Glu Ser Asp Val Ser Thr Ala Leu Arg Glu Lys Leu Arg	
885 890 895	
tct tcg ccg agg gcg gac gtg gac tcg acg acc ctc acg atg atc cgg	2736
Ser Ser Pro Arg Ala Asp Val Asp Ser Thr Thr Leu Thr Met Ile Arg	
900 905 910	
gca cag gca gcc gtg gtc ctc ggc cac tcc gat ccg aaa gag gtg gac	2784
Ala Gln Ala Ala Val Val Leu Gly His Ser Asp Pro Lys Glu Val Asp	
915 920 925	
ccg gat cgg acg ttc aag gac ctg ggc ttc gat tcc tcg atg gtg gtc	2832
Pro Asp Arg Thr Phe Lys Asp Leu Gly Phe Asp Ser Ser Met Val Val	
930 935 940	
gag ctg tgc gac cgc cta aac gcc gcc aca ggt ctg cga ctc gca ccg	2880
Glu Leu Cys Asp Arg Leu Asn Ala Ala Thr Gly Leu Arg Leu Ala Pro	
945 950 955 960	
agc gtc gtt ttc gac tgt cct acg ccg gac aag ctc gcc cgc cag gta	2928
Ser Val Val Phe Asp Cys Pro Thr Pro Asp Lys Leu Ala Arg Gln Val	
965 970 975	
cgg acg ttg ttg ttg ggc gag ccg gct ccc atg acg tca cac cgg ccg	2976
Arg Thr Leu Leu Leu Gly Glu Pro Ala Pro Met Thr Ser His Arg Pro	
980 985 990	
gac tcc gat gcg gac gag cct atc gcc gtg atc ggg atg ggc tgt cgg	3024
Asp Ser Asp Ala Asp Glu Pro Ile Ala Val Ile Gly Met Gly Cys Arg	
995 1000 1005	
ttt ccg ggt ggg gtg tcc tcg ccc gag gag ttg tgg cag ttg gtc gcc	3072
Phe Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Gln Leu Val Ala	
1010 1015 1020	
gct ggg cgg gac gtc gtg tcc gag ttc ccg gct gac cga ggt tgg gac	3120
Ala Gly Arg Asp Val Val Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp	
1025 1030 1035 1040	
ctg gag cgt gcg ggg aca tcg cac gtg cgc gcc ggc ggg ttc ttg cat	3168
Leu Glu Arg Ala Gly Thr Ser His Val Arg Ala Gly Gly Phe Leu His	
1045 1050 1055	
ggc gcc ccg gat ttt gac ccc ggg ttc ttc cgg att tcg ccg cgc gag	3216
Gly Ala Pro Asp Phe Asp Pro Gly Phe Phe Arg Ile Ser Pro Arg Glu	
1060 1065 1070	
gcg ttg gcg atg gat cca cag cag cgg ttg ctg ctg gaa atc gcc tgg	3264



Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ile Ala Trp	
1075	1080 1085
gaa gca gtc gaa cga ggc ggg atc aac ccg cag cac ctg cac gga agt	3312
Glu Ala Val Glu Arg Gly Gly Ile Asn Pro Gln His Leu His Gly Ser	
1090	1095 1100
caa acc ggg gtc ttc gtc ggc gcg acc tcc ctg gac tac ggg cca cgc	3360
Gln Thr Gly Val Phe Val Gly Ala Thr Ser Leu Asp Tyr Gly Pro Arg	
1105	1110 1115 1120
ctg cac gaa gcg tcc gag gag gcg gcc ggg tac gtg ctc acc ggc agc	3408
Leu His Glu Ala Ser Glu Glu Ala Ala Gly Tyr Val Leu Thr Gly Ser	
1125	1130 1135
acc acg agt gtg gcg tcg ggt cgg gtt gcg tat tcg ttc ggg ttc gag	3456
Thr Thr Ser Val Ala Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu	
1140	1145 1150
ggc cct gcg gtg acg gtg gat acg gcg tgt tcg tcg tcg ttg gtg gcc	3504
Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala	
1155	1160 1165
ctg cat ttg gcg tgt cag tcg ttg cgt tcg ggt gag tgt gat ctg gcg	3552
Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala	
1170	1175 1180
ttg gcc ggt ggt gtg acc gtg atg gcc acg ccg ggg atg ttc gtg gag	3600
Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Met Phe Val Glu	
1185	1190 1195 1200
ttt tcg cgg cag cgt ggt ttg gcg ccg gat ggg cgg tgc aag tcg ttc	3648
Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe	
1205	1210 1215
gcg gag gcc gcc gac ggc acc ggc tgg tcc gag ggt gct ggc ctg gtt	3696
Ala Glu Ala Ala Asp Gly Thr Gly Trp Ser Glu Gly Ala Gly Leu Val	
1220	1225 1230
cta ctg gag cgg ttg tcg gat gcc cgg cgg aat ggg cat gag gtg ctg	3744
Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Glu Val Leu	
1235	1240 1245
gcg gtt gtt cgt ggt agt gcg gtg aat cag gac ggt gcg tcg aat ggt	3792
Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly	
1250	1255 1260
ttg acc gcg ccg aat ggt tcg tcg cag cag ccg gtg att gcc cag gca	3840



Met Thr Tyr Leu Asp Glu Arg Pro Asp Val Ser Ser Leu Asp Val Gly	
1460 1465 1470	
tac tcg ctg gcg ttg aca cgg tcg gcg ctg gat gag cga gcg gtg gtg	4464
Tyr Ser Leu Ala Leu Thr Arg Ser Ala Leu Asp Glu Arg Ala Val Val	
1475 1480 1485	
ctg ggg tcg gac cgt gaa acg ttg ttg tgc ggt gtg aaa gcg ctg tct	4512
Leu Gly Ser Asp Arg Glu Thr Leu Leu Cys Gly Val Lys Ala Leu Ser	
1490 1495 1500	
gcc ggt cat gag gct tct ggg ttg gtg acc gga tct gtg ggg gct ggg	4560
Ala Gly His Glu Ala Ser Gly Leu Val Thr Gly Ser Val Gly Ala Gly	
1505 1510 1515 1520	
ggc cgc atc ggg ttt gtg ttt tcc ggt cag ggt ggt cag tgg ctg ggg	4608
Gly Arg Ile Gly Phe Val Phe Ser Gly Gln Gly Gly Gln Trp Leu Gly	
1525 1530 1535	
atg ggc cgg ggg ctt tac cgg gct ttt ccg gtg ttc gct gct gcc ttt	4656
Met Gly Arg Gly Leu Tyr Arg Ala Phe Pro Val Phe Ala Ala Ala Phe	
1540 1545 1550	
gac gaa gct tgt gcc gag ctg gat gcg cat ctg ggc cag gaa atc ggg	4704
Asp Glu Ala Cys Ala Glu Leu Asp Ala His Leu Gly Gln Glu Ile Gly	
1555 1560 1565	
gtt cgg gag gtg gtg tcc ggt tcg gat gcg cag ttg ctg gat cgg acg	4752
Val Arg Glu Val Val Ser Gly Ser Asp Ala Gln Leu Leu Asp Arg Thr	
1570 1575 1580	
ttg tgg gcg cag tcg ggt ttg ttc gcg ttg cag gtg ggc ttg ctg aag	4800
Leu Trp Ala Gln Ser Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Lys	
1585 1590 1595 1600	
ttg ctg gat tcg tgg ggg gtt cgg ccg agt gtg gtg ttg ggg cat tcg	4848
Leu Leu Asp Ser Trp Gly Val Arg Pro Ser Val Val Leu Gly His Ser	
1605 1610 1615	
gtg ggc gag ttg gcg gcg gcg ttc gcg gcg ggt gtg gtg tcg ttg tcg	4896
Val Gly Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Val Ser Leu Ser	
1620 1625 1630	
ggt gcg gct cgg ttg gtg gcg ggt cgt gcc cgg ttg atg cag gcg ttg	4944
Gly Ala Ala Arg Leu Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu	
1635 1640 1645	
ccg tct ggc ggt ggg atg ctg gcg gtg cct gct ggt gag gag ctg ttg	4992



His Val Arg Gly Gly Ala Val Asp Trp Arg Ser Phe Phe Ala Gly Thr	
1845	1850 1855
agg gcg aag caa atc gag ctg ccc acc tac gcc ttc caa cga cag cgg	5616
Arg Ala Lys Gln Ile Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg	
1860	1865 1870
tac tgg ctg aac gcg ctg cgt gaa tct tcc gcc ggc gac atg ggc agg	5664
Tyr Trp Leu Asn Ala Leu Arg Glu Ser Ser Ala Gly Asp Met Gly Arg	
1875	1880 1885
cgt gtc gaa gcg aag ttc tgg ggc gcc gtc gag cac gaa gat gtg gaa	5712
Arg Val Glu Ala Lys Phe Trp Gly Ala Val Glu His Glu Asp Val Glu	
1890	1895 1900
tcg ctt gca cgc gta ttg ggc att gtg gac gac ggc gct gct gtg gat	5760
Ser Leu Ala Arg Val Leu Gly Ile Val Asp Asp Gly Ala Ala Val Asp	
1905	1910 1915 1920
tcc ctg aga agc gcc ctt ccg gtg ttg gcc ggt tgg cag cga acc cgc	5808
Ser Leu Arg Ser Ala Leu Pro Val Leu Ala Gly Trp Gln Arg Thr Arg	
1925	1930 1935
acc acc gag tcc att atg gat cag cgg tgt tac cga att ggc tgg cgg	5856
Thr Thr Glu Ser Ile Met Asp Gln Arg Cys Tyr Arg Ile Gly Trp Arg	
1940	1945 1950
cag gta gcc gga ctc ccg ccg atg gga act gtt ttc ggt acc tgg ctg	5904
Gln Val Ala Gly Leu Pro Pro Met Gly Thr Val Phe Gly Thr Trp Leu	
1955	1960 1965
gtc ttc gcg cct cat ggc tgg tcc agc gaa ccg gag gtg gtg gac tgc	5952
Val Phe Ala Pro His Gly Trp Ser Ser Glu Pro Glu Val Val Asp Cys	
1970	1975 1980
gtt acg gca ctg cgg gca cgt ggt gcc tcg gtg gtg ttg gtg gaa gct	6000
Val Thr Ala Leu Arg Ala Arg Gly Ala Ser Val Val Leu Val Glu Ala	
1985	1990 1995 2000
gat ccc gac ccg acc tcc ttc ggc gac cgg gta cga acc ctg tgt tcg	6048
Asp Pro Asp Pro Thr Ser Phe Gly Asp Arg Val Arg Thr Leu Cys Ser	
2005	2010 2015
ggc ctt ccg gat ctt gtt ggc gtg ttg tca atg ttg tgc ttg gaa gaa	6096
Gly Leu Pro Asp Leu Val Gly Val Leu Ser Met Leu Cys Leu Glu Glu	
2020	2025 2030
tcg gtc ctt ccg gga ttt tct gcg gtg tca cgg ggt ttt gcg ttg acc	6144

Ser Val Leu Pro Gly Phe Ser Ala Val Ser Arg Gly Phe Ala Leu Thr  
 2035 2040 2045

gtg gag ttg gtg cgg gtt ttg cgg gca gct ggt gcg act gcc cgg ttg 6192  
 Val Glu Leu Val Arg Val Leu Arg Ala Ala Gly Ala Thr Ala Arg Leu  
 2050 2055 2060

tgg ttg ctg acg tgt ggt ggc gtg tcg gtg gga gat gta ccg gtt cgt 6240  
 Trp Leu Leu Thr Cys Gly Gly Val Ser Val Gly Asp Val Pro Val Arg  
 2065 2070 2075 2080

cca gcg cag gcc ctg gcg tgg ggg ttg ggg cgt gtt gtg ggg ttg gag 6288  
 Pro Ala Gln Ala Leu Ala Trp Gly Leu Gly Arg Val Val Gly Leu Glu  
 2085 2090 2095

cat ccg gac tgg tgg ggc ggc ttg atc gat att ccg gtc ttg ttc gac 6336  
 His Pro Asp Trp Trp Gly Gly Leu Ile Asp Ile Pro Val Leu Phe Asp  
 2100 2105 2110

gaa gac gct caa gag cgg ttg tcg att gtg ctg gca ggt ctc gat gag 6384  
 Glu Asp Ala Gln Glu Arg Leu Ser Ile Val Leu Ala Gly Leu Asp Glu  
 2115 2120 2125

gac gag gtc gcg atc cgt cct gac ggc atg ttc gcg cgt cgg ttg gta 6432  
 Asp Glu Val Ala Ile Arg Pro Asp Gly Met Phe Ala Arg Arg Leu Val  
 2130 2135 2140

cgc cac act gtc tca gct gat gtg aag aag gcg tgg cgc ccc agg gga 6480  
 Arg His Thr Val Ser Ala Asp Val Lys Lys Ala Trp Arg Pro Arg Gly  
 2145 2150 2155 2160

tcg gtg ctg gtg acg ggc ggc acg ggt ggt ttg ggg gcg cac gtt gct 6528  
 Ser Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala  
 2165 2170 2175

cgc tgg ctg gcc gac gcc gga gcc gaa cat gtg gcg atg gtg agt cga 6576  
 Arg Trp Leu Ala Asp Ala Gly Ala Glu His Val Ala Met Val Ser Arg  
 2180 2185 2190

cgc ggc gag cag gca ccg agt gct gag aag ttg cgg acg gaa ctg gag 6624  
 Arg Gly Glu Gln Ala Pro Ser Ala Glu Lys Leu Arg Thr Glu Leu Glu  
 2195 2200 2205

gat ctg ggt acc cgg gtg tcg atc gtg tca tgc gat gtg acc gat cgc 6672  
 Asp Leu Gly Thr Arg Val Ser Ile Val Ser Cys Asp Val Thr Asp Arg  
 2210 2215 2220

gag gcg ctc gcc gaa gtg ctg aaa gcc ctt ccg gct gaa aac ccg ttg 6720

Glu Ala Leu Ala Glu Val	Lys Ala Leu Pro Ala Glu Asn Pro Leu	
2225	2230	2235                  2240
acc gcg gta gtg cat gcg gca gcc gtg atc gag act ggt gat gcg gcg		6768
Thr Ala Val Val His Ala Ala Gly Val Ile Glu Thr Gly Asp Ala Ala		
2245	2250	2255
gca atg agc ctg gct gat ttc gat cac gtg ttg tcc gca aag gtg gcc		6816
Ala Met Ser Leu Ala Asp Phe Asp His Val Leu Ser Ala Lys Val Ala		
2260	2265	2270
ggg gcc gcg aat ctg gat gcc ttg ttg gcc gat gtg gaa ttg gac gcg		6864
Gly Ala Ala Asn Leu Asp Ala Leu Leu Ala Asp Val Glu Leu Asp Ala		
2275	2280	2285
tgc gtc ttg ttc tca tcg gtg tca gga gtt tgg gcc gct ggg gga cac		6912
Phe Val Leu Phe Ser Ser Val Ser Gly Val Trp Gly Ala Gly Gly His		
2290	2295	2300
ggg gct tac gca gcg gcg aat gcc tat ctg gat gcg ctc gcg gaa cag		6960
Gly Ala Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln		
2305	2310	2315                  2320
cgt cgg tcg cga ggg ctg gtc gcg act gcg gtg gcc tgg ggg ccg tgg		7008
Arg Arg Ser Arg Gly Leu Val Ala Thr Ala Val Ala Trp Gly Pro Trp		
2325	2330	2335
gcc gcc gag gcc atg gcc tcc gga gaa aca gga gac cag ctg cgc cga		7056
Ala Gly Glu Gly Met Ala Ser Gly Glu Thr Gly Asp Gln Leu Arg Arg		
2340	2345	2350
tac gcc ctt tcc cca atg gct ccg cag cac gcc atc gcc gga atc cgg		7104
Tyr Gly Leu Ser Pro Met Ala Pro Gln His Ala Ile Ala Gly Ile Arg		
2355	2360	2365
cag gcc gtg gaa cag gac gaa att tcc ctg gta gtg gcc gat gtc gat		7152
Gln Ala Val Glu Gln Asp Glu Ile Ser Leu Val Val Ala Asp Val Asp		
2370	2375	2380
tgg gca cgt ttc agc gcg gga ttg ctg gcg gct agg ccg cgg ccg ctg		7200
Trp Ala Arg Phe Ser Ala Gly Leu Leu Ala Ala Arg Pro Arg Pro Leu		
2385	2390	2395                  2400
ctg aac gaa ctg gcc gag gtc aag gaa ctc ctc gtc gat gcc cag ccc		7248
Leu Asn Glu Leu Ala Glu Val Lys Glu Leu Leu Val Asp Ala Gln Pro		
2405	2410	2415
gag gcg gga gtc ctt gcc gac gcg tcg ttg gaa tgg ccg cag cga ttg		7296

Glu Ala Gly Val Leu Ala Asp Ala Ser Leu Glu Trp Ala Gln Arg Leu	
2420	2425 2430
tcc gcg gca ccg agg ccg aca cag gaa cag ctg atc ctg gag ctg gta	7344
Ser Ala Ala Pro Arg Pro Thr Gln Glu Gln Leu Ile Leu Glu Leu Val	
2435	2440 2445
cgc ggc gaa acc gct ctg gtg ctg gga cac ccc ggg gca gcg gcc gtt	7392
Arg Gly Glu Thr Ala Leu Val Leu Gly His Pro Gly Ala Ala Val	
2450	2455 2460
gca tcg gaa cga gcc ttc aag gac agc gga ttc gac tcg cag gcc gcg	7440
Ala Ser Glu Arg Ala Phe Lys Asp Ser Gly Phe Asp Ser Gln Ala Ala	
2465	2470 2475 2480
gtc gaa ctc cgc gtt cgg ctc aat cga gct acc ggc ctc cag ttg cca	7488
Val Glu Leu Arg Val Arg Leu Asn Arg Ala Thr Gly Leu Gln Leu Pro	
2485	2490 2495
tcg aca att atc ttc agc cat ccc acg cct gcg gaa ctg gct gcg gag	7536
Ser Thr Ile Ile Phe Ser His Pro Thr Pro Ala Glu Leu Ala Ala Glu	
2500	2505 2510
ctg cgg gcg agg ctt ctt ccc gag tcc gca gga gca ggc att ccc gag	7584
Leu Arg Ala Arg Leu Leu Pro Glu Ser Ala Gly Ala Gly Ile Pro Glu	
2515	2520 2525
gag gac gag gcg cga atc aga gcg gca ctg acg tcg atc ccg ttc ccg	7632
Glu Asp Glu Ala Arg Ile Arg Ala Ala Leu Thr Ser Ile Pro Phe Pro	
2530	2535 2540
gcc ttg cgc gag gca ggc ttg gtg agt ccg ctg ctc gca ctt gcc gga	7680
Ala Leu Arg Glu Ala Gly Leu Val Ser Pro Leu Leu Ala Leu Ala Gly	
2545	2550 2555 2560
cac ccg gtc gac tcc ggt atc tcc tcg gac gat gcg gcc gcg acc tcg	7728
His Pro Val Asp Ser Gly Ile Ser Ser Asp Asp Ala Ala Ala Thr Ser	
2565	2570 2575
atc gat gcg atg gat gta gcc ggc ctc gtc gaa gca gcg ctg ggc gaa	7776
Ile Asp Ala Met Asp Val Ala Gly Leu Val Glu Ala Ala Leu Gly Glu	
2580	2585 2590
cgc gag tcc tga	7788
Arg Glu Ser	
2595	



<210> 42  
 <211> 2595  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 42  
 Met Ser Glu Ala Gly Asn Leu Ile Ala Val Ile Gly Leu Ser Cys Arg  
   1                  5                  10                  15  
 Leu Pro Gln Ala Pro Asp Pro Ala Ser Phe Trp Arg Leu Leu Arg Thr  
           20                  25                  30  
 Gly Thr Asp Ala Ile Thr Thr Val Pro Glu Gly Arg Trp Gly Asp Pro  
           35                  40                  45  
 Leu Pro Gly Arg Asp Ala Pro Lys Gly Pro Glu Trp Gly Gly Phe Leu  
       50                  55                  60  
 Ala Asp Val Asp Cys Phe Asp Pro Glu Phe Phe Gly Ile Ser Pro Arg  
   65                  70                  75                  80  
 Glu Ala Ala Ala Val Asp Pro Gln Gln Arg Leu Ala Leu Glu Leu Ala  
                   85                  90                  95  
 Trp Glu Ala Leu Glu Asp Ala Gly Ile Pro Ala Gly Glu Leu Arg Gly  
           100                  105                  110  
 Thr Ala Ala Gly Val Phe Met Gly Ala Ile Ser Asp Asp Tyr Ala Ala  
           115                  120                  125  
 Leu Leu Arg Glu Ser Pro Pro Glu Val Ala Ala Gln Tyr Arg Leu Thr  
       130                  135                  140  
 Gly Thr His Arg Ser Leu Ile Ala Asn Arg Val Ser Tyr Val Leu Gly  
 145                  150                  155                  160  
 Leu Arg Gly Pro Ser Leu Thr Val Asp Ser Gly Gln Ser Ser Ser Leu  
           165                  170                  175  
 Val Gly Val His Leu Ala Ser Glu Ser Leu Arg Arg Gly Glu Cys Thr  
           180                  185                  190  
 Ile Ala Leu Ala Gly Gly Val Asn Leu Asn Leu Ala Ala Glu Ser Asn  
           195                  200                  205  
 Ser Ala Leu Met Asp Phe Gly Ala Leu Ser Pro Asp Gly Arg Cys Phe  
       210                  215                  220

Thr Phe Asp Val Arg Ala Asn Gly Tyr Val Arg Gly Glu Gly Gly Gly  
 225 230 235 240

Leu Val Val Leu Lys Lys Ala Asp Gln Ala His Ala Asp Gly Asp Arg  
 245 250 255

Ile Tyr Cys Leu Ile Arg Gly Ser Ala Val Asn Asn Asp Gly Gly Gly  
 260 265 270

Ala Gly Leu Thr Val Pro Ala Ala Asp Ala Gln Ala Glu Leu Leu Arg  
 275 280 285

Gln Ala Tyr Arg Asn Ala Gly Val Asp Pro Ala Ala Val Gln Tyr Val  
 290 295 300

Glu Leu His Gly Ser Ala Thr Arg Val Gly Asp Pro Val Glu Ala Ala  
 305 310 315 320

Ala Leu Gly Ala Val Leu Gly Ala Ala Arg Arg Pro Gly Asp Glu Leu  
 325 330 335

Arg Val Gly Ser Ala Lys Thr Asn Val Gly His Leu Glu Ala Ala Ala  
 340 345 350

Gly Val Thr Gly Leu Leu Lys Thr Ala Leu Ser Ile Trp His Arg Glu  
 355 360 365

Leu Pro Pro Ser Leu His Phe Thr Ala Pro Asn Pro Glu Ile Pro Leu  
 370 375 380

Asp Glu Leu Asn Leu Arg Val Gln Arg Asp Leu Arg Pro Trp Pro Glu  
 385 390 395 400

Ser Glu Gly Pro Leu Leu Ala Gly Val Ser Ala Phe Gly Met Gly Gly  
 405 410 415

Thr Asn Cys His Leu Val Leu Ser Gly Thr Ser Arg Val Glu Arg Arg  
 420 425 430

Arg Ser Gly Pro Ala Glu Ala Thr Met Pro Trp Val Leu Ser Ala Arg  
 435 440 445

Thr Pro Val Ala Leu Arg Ala Gln Ala Ala Arg Leu His Thr His Leu  
 450 455 460

Asn Thr Ala Gly Gln Ser Pro Leu Asp Val Ala Tyr Ser Leu Ala Thr  
 465 470 475 480

Thr Arg Ser Ala Leu Pro His Arg Ala Ala Leu Val Ala Asp Asp Glu  
 485 490 495

Pro Lys Leu Leu Ala Gly Leu Lys Ala Leu Ala Asp Gly Asp Asp Ala  
 500 505 510

Pro Thr Leu Cys His Gly Ala Thr Ser Gly Glu Arg Ala Ala Val Phe  
 515 520 525

Val Phe Pro Gly Gln Gly Ser Gln Trp Ile Gly Met Gly Arg Gln Leu  
 530 535 540

Leu Glu Thr Ser Glu Val Phe Ala Ala Ser Met Ser Asp Cys Ala Asp  
 545 550 555 560

Ala Leu Ala Pro His Leu Asp Trp Ser Leu Leu Asp Val Leu Arg Asn  
 565 570 575

Ala Ala Gly Ala Ala His Leu Asp His Asp Asp Val Val Gln Pro Ala  
 580 585 590

Leu Phe Ala Ile Met Val Ser Leu Ala Glu Leu Trp Arg Ser Trp Gly  
 595 600 605

Val Arg Pro Val Ala Val Val Gly His Ser Gln Gly Glu Ile Ala Ala  
 610 615 620

Ala Cys Val Ala Gly Ala Leu Ser Val Arg Asp Ala Ala Arg Val Val  
 625 630 635 640

Ala Val Arg Ser Arg Leu Leu Thr Ala Leu Ala Gly Ser Gly Ala Met  
 645 650 655

Ala Ser Leu Gln His Pro Ala Glu Glu Val Arg Gln Ile Leu Leu Pro  
 660 665 670

Trp Arg Asp Arg Ile Gly Val Ala Gly Val Asn Gly Pro Ser Ser Thr  
 675 680 685

Leu Val Ser Gly Asp Arg Glu Ala Met Ala Glu Leu Leu Ala Glu Cys  
 690 695 700

Ala Asp Arg Glu Leu Arg Met Arg Arg Ile Pro Val Glu Tyr Ala Ser  
 705 710 715 720

His Ser Pro His Ile Glu Val Val Arg Asp Glu Leu Leu Gly Leu Leu  
 725 730 735

Ala Pro Val Glu Pro Arg Thr Gly Ser Ile Pro Ile Tyr Ser Thr Thr  
 740 745 750

Thr Gly Asp Leu Leu Asp Arg Pro Met Asp Ala Asp Tyr Trp Tyr Arg  
 755 760 765

Asn Leu Arg Gln Pro Val Leu Phe Glu Ala Ala Val Glu Ala Leu Leu  
 770 775 780

Lys Arg Gly Tyr Asp Ala Phe Ile Glu Ile Ser Pro His Pro Val Leu  
 785 790 795 800

Thr Ala Asn Ile Gln Glu Thr Ala Val Arg Ala Gly Arg Glu Val Val  
 805 810 815

Ala Leu Gly Thr Leu Arg Arg Gly Glu Gly Gly Met Arg Gln Ala Leu  
 820 825 830

Thr Ser Leu Ala Arg Ala His Val His Gly Val Ala Ala Asp Trp His  
 835 840 845

Ala Val Phe Ala Gly Thr Gly Ala Gln Arg Val Asp Leu Pro Thr Tyr  
 850 855 860

Ala Phe Gln Arg Gln Arg Tyr Trp Leu Asp Ala Lys Leu Pro Asp Val  
 865 870 875 880

Ala Met Pro Glu Ser Asp Val Ser Thr Ala Leu Arg Glu Lys Leu Arg  
 885 890 895

Ser Ser Pro Arg Ala Asp Val Asp Ser Thr Thr Leu Thr Met Ile Arg  
 900 905 910

Ala Gln Ala Ala Val Val Leu Gly His Ser Asp Pro Lys Glu Val Asp  
 915 920 925

Pro Asp Arg Thr Phe Lys Asp Leu Gly Phe Asp Ser Ser Met Val Val  
 930 935 940

Glu Leu Cys Asp Arg Leu Asn Ala Ala Thr Gly Leu Arg Leu Ala Pro  
 945 950 955 960

Ser Val Val Phe Asp Cys Pro Thr Pro Asp Lys Leu Ala Arg Gln Val  
 965 970 975

Arg Thr Leu Leu Leu Gly Glu Pro Ala Pro Met Thr Ser His Arg Pro  
 980 985 990

Asp Ser Asp Ala Asp Glu Pro Ile Ala Val Ile Gly Met Gly Cys Arg  
995 1000 1005

Phe Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Gln Leu Val Ala  
1010 1015 1020

Ala Gly Arg Asp Val Val Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp  
025 1030 1035 1040

Leu Glu Arg Ala Gly Thr Ser His Val Arg Ala Gly Gly Phe Leu His  
1045 1050 1055

Gly Ala Pro Asp Phe Asp Pro Gly Phe Phe Arg Ile Ser Pro Arg Glu  
1060 1065 1070

Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ile Ala Trp  
1075 1080 1085

Glu Ala Val Glu Arg Gly Gly Ile Asn Pro Gln His Leu His Gly Ser  
1090 1095 1100

Gln Thr Gly Val Phe Val Gly Ala Thr Ser Leu Asp Tyr Gly Pro Arg  
105 1110 1115 1120

Leu His Glu Ala Ser Glu Glu Ala Ala Gly Tyr Val Leu Thr Gly Ser  
1125 1130 1135

Thr Thr Ser Val Ala Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu  
1140 1145 1150

Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala  
1155 1160 1165

Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala  
1170 1175 1180

Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Met Phe Val Glu  
185 1190 1195 1200

Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe  
1205 1210 1215

Ala Glu Ala Ala Asp Gly Thr Gly Trp Ser Glu Gly Ala Gly Leu Val  
1220 1225 1230

Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Glu Val Leu  
1235 1240 1245

Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly  
1250 1255 1260

Leu Thr Ala Pro Asn Gly Ser Ser Gln Gln Arg Val Ile Ala Gln Ala  
265 1270 1275 1280

Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp Ala Val Glu Ala  
1285 1290 1295

His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu  
1300 1305 1310

Ile Ala Thr Tyr Gly Gln Gly Arg Leu Pro Glu Arg Pro Leu Trp Leu  
1315 1320 1325

Gly Ser Met Lys Ser Asn Ile Gly His Ala Gln Ala Ala Ala Gly Ile  
1330 1335 1340

Ala Gly Val Met Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro  
345 1350 1355 1360

Arg Thr Leu His Val Asp Glu Pro Thr Ser Gly Val Asp Trp Ser Ala  
1365 1370 1375

Gly Thr Val Gln Leu Leu Thr Glu Asn Thr Pro Trp Pro Gly Ser Gly  
1380 1385 1390

Arg Val Arg Arg Val Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn  
1395 1400 1405

Ala His Val Ile Leu Glu Gln Pro Pro Gly Val Pro Ser Gln Ser Ala  
1410 1415 1420

Gly Pro Gly Ser Gly Ser Val Val Asp Val Pro Val Val Pro Trp Met  
425 1430 1435 1440

Val Ser Gly Lys Thr Pro Glu Ala Leu Ser Ala Gln Ala Thr Ala Leu  
1445 1450 1455

Met Thr Tyr Leu Asp Glu Arg Pro Asp Val Ser Ser Leu Asp Val Gly  
1460 1465 1470

Tyr Ser Leu Ala Leu Thr Arg Ser Ala Leu Asp Glu Arg Ala Val Val  
1475 1480 1485

Leu Gly Ser Asp Arg Glu Thr Leu Leu Cys Gly Val Lys Ala Leu Ser  
1490 1495 1500

Ala Gly His Glu Ala Ser Gly Leu Val Thr Gly Ser Val Gly Ala Gly  
 505 1510 1515 1520

Gly Arg Ile Gly Phe Val Phe Ser Gly Gln Gly Gly Gln Trp Leu Gly  
 1525 1530 1535

Met Gly Arg Gly Leu Tyr Arg Ala Phe Pro Val Phe Ala Ala Ala Phe  
 1540 1545 1550

Asp Glu Ala Cys Ala Glu Leu Asp Ala His Leu Gly Gln Glu Ile Gly  
 1555 1560 1565

Val Arg Glu Val Val Ser Gly Ser Asp Ala Gln Leu Leu Asp Arg Thr  
 1570 1575 1580

Leu Trp Ala Gln Ser Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Lys  
 585 1590 1595 1600

Leu Leu Asp Ser Trp Gly Val Arg Pro Ser Val Val Leu Gly His Ser  
 1605 1610 1615

Val Gly Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Val Ser Leu Ser  
 1620 1625 1630

Gly Ala Ala Arg Leu Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu  
 1635 1640 1645

Pro Ser Gly Gly Gly Met Leu Ala Val Pro Ala Gly Glu Glu Leu Leu  
 1650 1655 1660

Trp Ser Leu Leu Ala Asp Gln Gly Asp Arg Val Gly Ile Ala Ala Val  
 665 1670 1675 1680

Asn Ala Ala Gly Ser Val Val Leu Ser Gly Asp Arg Asp Val Leu Asp  
 1685 1690 1695

Asp Leu Ala Gly Arg Leu Asp Gly Gln Gly Ile Arg Ser Arg Trp Leu  
 1700 1705 1710

Arg Val Ser His Ala Phe His Ser Tyr Arg Met Asp Pro Met Leu Ala  
 1715 1720 1725

Glu Phe Ala Glu Leu Ala Arg Thr Val Asp Tyr Arg Arg Cys Glu Val  
 1730 1735 1740

Pro Ile Val Ser Thr Leu Thr Gly Asp Leu Asp Asp Ala Gly Arg Met  
 745 1750 1755 1760

Ser Gly Pro Asp Trp Val Arg Gln Val Arg Glu Pro Val Arg Phe  
 1765 1770 1775

Ala Asp Gly Val Gln Ala Leu Val Glu His Asp Val Ala Thr Val Val  
 1780 1785 1790

Glu Leu Gly Pro Asp Gly Ala Leu Ser Ala Leu Ile Gln Glu Cys Val  
 1795 1800 1805

Ala Ala Ser Asp His Ala Gly Arg Leu Ser Ala Val Pro Ala Met Arg  
 1810 1815 1820

Arg Asn Gln Asp Glu Ala Gln Lys Val Met Thr Ala Leu Ala His Val  
 1825 1830 1835 1840

His Val Arg Gly Gly Ala Val Asp Trp Arg Ser Phe Phe Ala Gly Thr  
 1845 1850 1855

Arg Ala Lys Gln Ile Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg  
 1860 1865 1870

Tyr Trp Leu Asn Ala Leu Arg Glu Ser Ser Ala Gly Asp Met Gly Arg  
 1875 1880 1885

Arg Val Glu Ala Lys Phe Trp Gly Ala Val Glu His Glu Asp Val Glu  
 1890 1895 1900

Ser Leu Ala Arg Val Leu Gly Ile Val Asp Asp Gly Ala Ala Val Asp  
 1905 1910 1915 1920

Ser Leu Arg Ser Ala Leu Pro Val Leu Ala Gly Trp Gln Arg Thr Arg  
 1925 1930 1935

Thr Thr Glu Ser Ile Met Asp Gln Arg Cys Tyr Arg Ile Gly Trp Arg  
 1940 1945 1950

Gln Val Ala Gly Leu Pro Pro Met Gly Thr Val Phe Gly Thr Trp Leu  
 1955 1960 1965

Val Phe Ala Pro His Gly Trp Ser Ser Glu Pro Glu Val Val Asp Cys  
 1970 1975 1980

Val Thr Ala Leu Arg Ala Arg Gly Ala Ser Val Val Leu Val Glu Ala  
 1985 1990 1995 2000

Asp Pro Asp Pro Thr Ser Phe Gly Asp Arg Val Arg Thr Leu Cys Ser  
 2005 2010 2015



Gly Leu Pro Asp Leu Val Gly Val Leu Ser Met Leu Cys Leu Glu Glu  
 2020 2025 2030

Ser Val Leu Pro Gly Phe Ser Ala Val Ser Arg Gly Phe Ala Leu Thr  
 2035 2040 2045

Val Glu Leu Val Arg Val Leu Arg Ala Ala Gly Ala Thr Ala Arg Leu  
 2050 2055 2060

Trp Leu Leu Thr Cys Gly Gly Val Ser Val Gly Asp Val Pro Val Arg  
 065 2070 2075 2080

Pro Ala Gln Ala Leu Ala Trp Gly Leu Gly Arg Val Val Gly Leu Glu  
 2085 2090 2095

His Pro Asp Trp Trp Gly Gly Leu Ile Asp Ile Pro Val Leu Phe Asp  
 2100 2105 2110

Glu Asp Ala Gln Glu Arg Leu Ser Ile Val Leu Ala Gly Leu Asp Glu  
 2115 2120 2125

Asp Glu Val Ala Ile Arg Pro Asp Gly Met Phe Ala Arg Arg Leu Val  
 2130 2135 2140

Arg His Thr Val Ser Ala Asp Val Lys Lys Ala Trp Arg Pro Arg Gly  
 145 2150 2155 2160

Ser Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala  
 2165 2170 2175

Arg Trp Leu Ala Asp Ala Gly Ala Glu His Val Ala Met Val Ser Arg  
 2180 2185 2190

Arg Gly Glu Gln Ala Pro Ser Ala Glu Lys Leu Arg Thr Glu Leu Glu  
 2195 2200 2205

Asp Leu Gly Thr Arg Val Ser Ile Val Ser Cys Asp Val Thr Asp Arg  
 2210 2215 2220

Glu Ala Leu Ala Glu Val Leu Lys Ala Leu Pro Ala Glu Asn Pro Leu  
 225 2230 2235 2240

Thr Ala Val Val His Ala Ala Gly Val Ile Glu Thr Gly Asp Ala Ala  
 2245 2250 2255

Ala Met Ser Leu Ala Asp Phe Asp His Val Leu Ser Ala Lys Val Ala  
 2260 2265 2270

Gly Ala Ala Asn Leu Asp Ala Leu Leu Ala Asp Val Glu Leu Asp Ala  
 2275 2280 2285

Phe Val Leu Phe Ser Ser Val Ser Gly Val Trp Gly Ala Gly Gly His  
 2290 2295 2300

Gly Ala Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln  
 305 2310 2315 2320

Arg Arg Ser Arg Gly Leu Val Ala Thr Ala Val Ala Trp Gly Pro Trp  
 2325 2330 2335

Ala Gly Glu Gly Met Ala Ser Gly Glu Thr Gly Asp Gln Leu Arg Arg  
 2340 2345 2350

Tyr Gly Leu Ser Pro Met Ala Pro Gln His Ala Ile Ala Gly Ile Arg  
 2355 2360 2365

Gln Ala Val Glu Gln Asp Glu Ile Ser Leu Val Val Ala Asp Val Asp  
 2370 2375 2380

Trp Ala Arg Phe Ser Ala Gly Leu Leu Ala Ala Arg Pro Arg Pro Leu  
 385 2390 2395 2400

Leu Asn Glu Leu Ala Glu Val Lys Glu Leu Leu Val Asp Ala Gln Pro  
 2405 2410 2415

Glu Ala Gly Val Leu Ala Asp Ala Ser Leu Glu Trp Arg Gln Arg Leu  
 2420 2425 2430

Ser Ala Ala Pro Arg Pro Thr Gln Glu Gln Leu Ile Leu Glu Leu Val  
 2435 2440 2445

Arg Gly Glu Thr Ala Leu Val Leu Gly His Pro Gly Ala Ala Ala Val  
 2450 2455 2460

Ala Ser Glu Arg Ala Phe Lys Asp Ser Gly Phe Asp Ser Gln Ala Ala  
 465 2470 2475 2480

Val Glu Leu Arg Val Arg Leu Asn Arg Ala Thr Gly Leu Gln Leu Pro  
 2485 2490 2495

Ser Thr Ile Ile Phe Ser His Pro Thr Pro Ala Glu Leu Ala Ala Glu  
 2500 2505 2510

Leu Arg Ala Arg Leu Leu Pro Glu Ser Ala Gly Ala Gly Ile Pro Glu  
 2515 2520 2525

Glu Asp Glu Ala Arg Ile Arg Ala Ala Leu Thr Ser Ile Pro Phe Pro  
 2530 2535 2540

Ala Leu Arg Glu Ala Gly Leu Val Ser Pro Leu Leu Ala Leu Ala Gly  
 545 2550 2555 2560

His Pro Val Asp Ser Gly Ile Ser Ser Asp Asp Ala Ala Ala Thr Ser  
 2565 2570 2575

Ile Asp Ala Met Asp Val Ala Gly Leu Val Glu Ala Ala Leu Gly Glu  
 2580 2585 2590

Arg Glu Ser  
 2595

<210> 43  
 <211> 6459  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(6456)  
 <223> ORF19; Polyketidsynthase

<400> 43  
 atg acg gtg acc acc agt tac gaa gaa gtt gtc gag gca ctg cga gca 48  
 Met Thr Val Thr Thr Ser Tyr Glu Glu Val Val Glu Ala Leu Arg Ala  
 1 5 10 15  
 tcg ctc aag gag aac gaa cgc ctc cgg cgc ggc agg gat cgg ttc tcc 96  
 Ser Leu Lys Glu Asn Glu Arg Leu Arg Arg Gly Arg Asp Arg Phe Ser  
 20 25 30  
 gcg gag aag gac gat ccc atc gcg atc gtg gcg atg agt tgt cgt tat 144  
 Ala Glu Lys Asp Asp Pro Ile Ala Ile Val Ala Met Ser Cys Arg Tyr  
 35 40 45  
 ccc ggt cag gtc tcc tcg ccg gag gac ctg tgg caa ctg gct gcc ggc 192  
 Pro Gly Gln Val Ser Ser Pro Glu Asp Leu Trp Gln Leu Ala Ala Gly  
 50 55 60  
 ggt gtg gac gcg atc tcc gaa gtt ccg ggg gat cgc gga tgg gac ctg 240  
 Gly Val Asp Ala Ile Ser Glu Val Pro Gly Asp Arg Gly Trp Asp Leu  
 65 70 75 80

gat ggc gtg ttc ggc ccg gac tcc gat cgt cct ggc acc tcg tat gcc	288
Asp Gly Val Phe Val Pro Asp Ser Asp Arg Pro Gly Thr Ser Tyr Ala	
85 90 95	
tgc gcg ggc ggt ttt ctt cag ggc gtg tcg gag ttc gac gcg ggt ttc	336
Cys Ala Gly Gly Phe Leu Gln Gly Val Ser Glu Phe Asp Ala Gly Phe	
100 105 110	
ttc ggg att tcg ccg cgt gag gcg ctg gcg atg gat ccg cag cag cgg	384
Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg	
115 120 125	
ttg ctg ctg gaa gtc gcg tgg gag gtc ttc gag cgg gct ggg ctg gag	432
Leu Leu Leu Glu Val Ala Trp Glu Val Phe Glu Arg Ala Gly Leu Glu	
130 135 140	
cag cgg tcg aca cgc ggt tcc cgc gtt ggc gtg ttc gtc ggc acc aat	480
Gln Arg Ser Thr Arg Gly Ser Arg Val Gly Val Phe Val Gly Thr Asn	
145 150 155 160	
ggc cag gac tac gcg tcg tgg ttg cgg acg ccg ccg cct gcg gtg gca	528
Gly Gln Asp Tyr Ala Ser Trp Leu Arg Thr Pro Pro Pro Ala Val Ala	
165 170 175	
ggt cat gtg ctg acg ggc ggt gcg gca gcg gtt ctt tcg ggc cgg gtt	576
Gly His Val Leu Thr Gly Gly Ala Ala Ala Val Leu Ser Gly Arg Val	
180 185 190	
gcg tat tcg ttc ggg ttc gag ggt cct gcg gtg acg gtg gat acg gcg	624
Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala	
195 200 205	
tgt tcg tcg tcg ttg gtg gcg ttg cac ctg gcg ggg caa gca ctg cgg	672
Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu Arg	
210 215 220	
gcc ggt gag tgc gac ctt gcc ctt gcc ggt ggc gtc acg gtg atg tcg	720
Ala Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser	
225 230 235 240	
acg ccg aag gtg ttc ctg gag ttc tcc cgc caa cgg ggt ctc gcg ccg	768
Thr Pro Lys Val Phe Leu Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro	
245 250 255	
gat ggg cgg tgc aag tcg ttc gcg gcg ggt gcg gat ggc act gga tgg	816
Asp Gly Arg Cys Lys Ser Phe Ala Ala Gly Ala Asp Gly Thr Gly Trp	
260 265 270	

ggt gag ggt gcc gga ctg ttg ttg ctg gag cgg ttg tcg gat gcc cgg	964
Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg	
275 280 285	
cgg aat ggg cat gag gtg ctg gcg gtt gtt cgt ggt agt gcg gtg aat	912
Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn	
290 295 300	
cag gac ggt gcg tcg aat ggt ttg acc gcg ccg aat ggt tcg tcg cag	960
Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Ser Ser Gln	
305 310 315 320	
cag cgg gtg att acc cag gcg ttg gcg agt gcg ggg ttg tcg gtg tcc	1008
Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser	
325 330 335	
gat gtg gat gct gtg gag gcg cat ggg acg ggc acg cgg ctt ggt gat	1056
Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp	
340 345 350	
ccg atc gag gcg cag gcg ctg atc gcc acc tac ggc cgt gat cgt gat	1104
Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg Asp	
355 360 365	
cct ggc cgg ccg ttg tgg ttg ggg tcg gtc aag tcg aac atc ggt cat	1152
Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His	
370 375 380	
acg caa gcg gcg gcg ggt gtg gct ggt gtg atc aag atg gtg atg gcg	1200
Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala	
385 390 395 400	
atg cgg cac ggg cag ctg cca cgc acg ttg cac gtg gaa tcg ccg tcg	1248
Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Glu Ser Pro Ser	
405 410 415	
ccg gag gtg gat tgg tcg gcg ggg acg gtt caa ctc ctt acg gag aac	1296
Pro Glu Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu Asn	
420 425 430	
acg ccc tgg ccc agg agt ggt cgt gtt cgt cgg gtg ggg gtg tcg tcg	1344
Thr Pro Trp Pro Arg Ser Gly Arg Val Arg Arg Val Gly Val Ser Ser	
435 440 445	
ttc ggg atc agt ggt act aac gcg cac gtc atc ctc gaa cag ccc ccg	1392
Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro Pro	
450 455 460	

gga gtg ccg agt c	ct gcg ggg ccg ggt tcg ggt t	gtc gtg gat	1440
Gly Val Pro Ser Gln Ser Ala Gly Pro Gly Ser Gly Ser Val Val Asp			
465	470	475	480
ggt ccg gtg gtg ccg tgg atg gtg tcg ggc aaa aca ccc gaa gcg cta			1488
Val Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala Leu			
485	490	495	
tcc gcg cag gca acg gcg ttg atg acc tat ctg gac gag cga cct gat			1536
Ser Ala Gln Ala Thr Ala Leu Met Thr Tyr Leu Asp Glu Arg Pro Asp			
500	505	510	
gtc tcc tcg ctg gat gtt ggg tac tcg ctg gcg ttg aca cgg tcg gcg			1584
Val Ser Ser Leu Asp Val Gly Tyr Ser Leu Ala Leu Thr Arg Ser Ala			
515	520	525	
ctg gat gag cga gcg gtg gtg ctg ggg tcg gac cgt gaa acg ttg ttg			1632
Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Glu Thr Leu Leu			
530	535	540	
tgc ggt gtg aaa gcg ctg tct gcc ggt cat gag gct tct ggg ttg gtg			1680
Cys Gly Val Lys Ala Leu Ser Ala Gly His Glu Ala Ser Gly Leu Val			
545	550	555	560
acc gga tct gtg ggg gct ggg ggc cgc atc ggg ttt gtg ttt tcc ggt			1728
Thr Gly Ser Val Gly Ala Gly Gly Arg Ile Gly Phe Val Phe Ser Gly			
565	570	575	
cag ggt ggt cag tgg ctg ggg atg ggc cgg ggg ctt tac cgg gct ttt			1776
Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu Tyr Arg Ala Phe			
580	585	590	
ccg gtg ttc gct gct gcc ttt gac gaa gct tgt gcc gag ctg gat gca			1824
Pro Val Phe Ala Ala Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala			
595	600	605	
cat ctg ggc cag gaa atc ggg gtt cgg gag gtg gtg tcc ggt tcg gat			1872
His Leu Gly Gln Glu Ile Gly Val Arg Glu Val Val Ser Gly Ser Asp			
610	615	620	
gcg cag ttg ctg gat cgg acg ttg tgg gcg cag tcg ggt ttg ttc gcg			1920
Ala Gln Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Phe Ala			
625	630	635	640
ttg cag gtg ggc ttg ctg aag ttg ctg gat tcg tgg ggg gtt cgg ccg			1968
Leu Gln Val Gly Leu Leu Lys Leu Leu Asp Ser Trp Gly Val Arg Pro			
645	650	655	

agt gtg gtg ttg ggg cat tcg gtg ggc gag ttg gcg gcg gcg ttc gcg	2016
Ser Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala	
660 665 670	
gcg ggt gtg gtg tcg ttg tcg ggt gcg gct cgg ttg gtg gcg ggt cgt	2064
Ala Gly Val Val Ser Leu Ser Gly Ala Ala Arg Leu Val Ala Gly Arg	
675 680 685	
gcc cgg ttg atg cag gcg ttg ccg tct ggc ggt ggg atg ctg gcg gtg	2112
Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Met Leu Ala Val	
690 695 700	
cct gct ggt gag gag ctg ttg tgg tcg ttg ttg gcc gat cag ggt gat	2160
Pro Ala Gly Glu Glu Leu Leu Trp Ser Leu Leu Ala Asp Gln Gly Asp	
705 710 715 720	
cgt gtg ggg atc gcc gcg gtc aac gct gcg ggg tcg gtg gtg ctc tct	2208
Arg Val Gly Ile Ala Ala Val Asn Ala Ala Gly Ser Val Val Leu Ser	
725 730 735	
ggt gat cgg gat gtg ctc gat gac ctt gcc ggt cgg ctg gac ggg caa	2256
Gly Asp Arg Asp Val Leu Asp Asp Leu Ala Gly Arg Leu Asp Gly Gln	
740 745 750	
ggg atc cgg tcg agg tgg ttg cgg gtg tcg cat gcg ttt cat tcg tat	2304
Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser Tyr	
755 760 765	
cgg atg gat ccg atg ctg gcg gag ttc gcc gaa ttg gca cga acc gtg	2352
Arg Met Asp Pro Met Leu Ala Glu Phe Ala Glu Leu Ala Arg Thr Val	
770 775 780	
gat tac cgg cgt tgt gaa gtg ccg atc gtg tcg acc ttg acc gga gac	2400
Asp Tyr Arg Arg Cys Glu Val Pro Ile Val Ser Thr Leu Thr Gly Asp	
785 790 795 800	
ctc gat gac gct ggc agg atg agc ggg ccc gac tac tgg gtg cgt cag	2448
Leu Asp Asp Ala Gly Arg Met Ser Gly Pro Asp Tyr Trp Val Arg Gln	
805 810 815	
gtg cga gag ccg gtc cgc ttc gcc gac ggt gtc cag gcg ctg gtc gag	2496
Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Gln Ala Leu Val Glu	
820 825 830	
cac gat gtg gcc act gtt gtc gag ctc ggt ccg gac ggg gcg ttg tcg	2544
His Asp Val Ala Thr Val Val Glu Leu Gly Pro Asp Gly Ala Leu Ser	
835 840 845	

gcg ctg atc cag gaa tgt gtc gcc gca tcc gat cac gcc ggg cgg ctg	2592
Ala Leu Ile Gln Glu Cys Val Ala Ala Ser Asp His Ala Gly Arg Leu	
850 855 860	
agc gcg gtc ccg gcg atg cgc agg aac cag gac gag gcg cag aag gtg	2640
Ser Ala Val Pro Ala Met Arg Arg Asn Gln Asp Glu Ala Gln Lys Val	
865 870 875 380	
atg acg gcc ctg gca cac gtc cac gta cgt ggt ggt gcg gtg gac tgg	2688
Met Thr Ala Leu Ala His Val His Val Arg Gly Gly Ala Val Asp Trp	
885 890 895	
cgg tcg ttc ttc gcc ggt acg gga gcg aaa caa atc gag ctg ccc acc	2736
Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Ile Glu Leu Pro Thr	
900 905 910	
tac gcc ttc caa cga cag cgg tac tgg ctg gtg cca tcg gat tcc ggt	2784
Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Val Pro Ser Asp Ser Gly	
915 920 925	
gat gtg aca ggt gcc ggt ctg gcc ggg gcg gag cat ccg ctg ttg ggt	2832
Asp Val Thr Gly Ala Gly Leu Ala Gly Ala Glu His Pro Leu Leu Gly	
930 935 940	
gct gtg gtg ccg gtc gcg ggt ggt gac gag gtg ttg ctg acc ggc agg	2880
Ala Val Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg	
945 950 955 960	
att tcg gtg cgg acg cat ccg tgg ctg gcc gaa cac cgg gtg ctg ggt	2928
Ile Ser Val Arg Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly	
965 970 975	
gaa gtg atc gtt gcg ggc acc gcg ttg ctg gag atc gcc ttg cac gcg	2976
Glu Val Ile Val Ala Gly Thr Ala Leu Leu Glu Ile Ala Leu His Ala	
980 985 990	
ggg gaa cgt ctt ggt tgt gaa cgg gtg gaa gag ctc acc ctg gaa gca	3024
Gly Glu Arg Leu Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala	
995 1000 1005	
ccg ctg gtc ctg ccg gag cgc ggg gcg atc cag gtt cag ctg cga gtg	3072
Pro Leu Val Leu Pro Glu Arg Gly Ala Ile Gln Val Gln Leu Arg Val	
1010 1015 1020	
ggc gcg ccc gag aat tcc gga cgc agg ccg atg gcg ctg tat tca cgc	3120
Gly Ala Pro Glu Asn Ser Gly Arg Arg Pro Met Ala Leu Tyr Ser Arg	
1025 1030 1035 1040	



ccc gaa ggg gcg gcg gag cat gac tgg acg cgg cac gcc acg ggc cgg 3168  
Pro Glu Gly Ala Ala Glu His Asp Trp Thr Arg His Ala Thr Gly Arg  
1045 1050 1055

ttg gcg cca ggc cgc ggc gag gcg gct gga gac ctg gcc gac tgg ccg 3216  
Leu Ala Pro Gly Arg Gly Glu Ala Ala Gly Asp Leu Ala Asp Trp Pro  
1060 1065 1070

gct cct ggc gcg ctg ccg gtc gac ctc gac gaa ttc tat cgg gac ctc 3264  
Ala Pro Gly Ala Leu Pro Val Asp Leu Asp Glu Phe Tyr Arg Asp Leu  
1075 1080 1085

gca gag ctt ggg ctg gag tac ggc ccg atc ttc caa ggg ctc aag gcg 3312  
Ala Glu Leu Gly Leu Glu Tyr Gly Pro Ile Phe Gln Gly Leu Lys Ala  
1090 1095 1100

gcc tgg cgg caa ggg gac gag gtg tac gcc gaa gcc gcg ctg ccg gga 3360  
Ala Trp Arg Gln Gly Asp Glu Val Tyr Ala Glu Ala Ala Leu Pro Gly  
1105 1110 1115 1120

acg gaa gat tct ggt ttc ggg gtg cat ccg gca ctg ctg gac gcg gct 3408  
Thr Glu Asp Ser Gly Phe Gly Val His Pro Ala Leu Leu Asp Ala Ala  
1125 1130 1135

ctg cac gca acg gct gtc cga gac atg gat gac gca cgc ttg ccg ttc 3456  
Leu His Ala Thr Ala Val Arg Asp Met Asp Asp Ala Arg Leu Pro Phe  
1140 1145 1150

cag tgg gaa ggt gtg tcc ctg cac gcc aag gcc gcg ccg gct ttg cgg 3504  
Gln Trp Glu Gly Val Ser Leu His Ala Lys Ala Ala Pro Ala Leu Arg  
1155 1160 1165

gtc cgc gtg gtc ccg gct ggt gac gat gcc aag tcc ctg ctg gtt tgt 3552  
Val Arg Val Val Pro Ala Gly Asp Asp Ala Lys Ser Leu Leu Val Cys  
1170 1175 1180

gat ggc acc ggt cga ccg gtg atc tcg gtg gac cga ctc gta ttg cgg 3600  
Asp Gly Thr Gly Arg Pro Val Ile Ser Val Asp Arg Leu Val Leu Arg  
1185 1190 1195 1200

tcg gct gcg gcc cgg cgg acc ggt gcg cgc cga cag gcc cat caa gct 3648  
Ser Ala Ala Ala Arg Arg Thr Gly Ala Arg Arg Gln Ala His Gln Ala  
1205 1210 1215

cgg ttg tac cgg ttg agc tgg cca acg gtt caa ctg ccg aca tcc gct 3696  
Arg Leu Tyr Arg Leu Ser Trp Pro Thr Val Gln Leu Pro Thr Ser Ala  
1220 1225 1230

cag cca ccg tcc tgc gtg ctt ctc ggc acc tca gaa gtg tcc gct gac Gln Pro Pro Ser Cys Val Leu Leu Gly Thr Ser Glu Val Ser Ala Asp 1235 1240 1245	3744
ata cag gtg tat ccg gac ctc cgg tcg ttg acg gct gcg ttg gat gcc Ile Gln Val Tyr Pro Asp Leu Arg Ser Leu Thr Ala Ala Leu Asp Ala 1250 1255 1260	3792
ggt gcc gaa cca ccc ggc gtc gtc atc gca ccc acg ccc ccc ggc ggt Gly Ala Glu Pro Pro Gly Val Val Ile Ala Pro Thr Pro Pro Gly Gly 1265 1270 1275 1280	3840
gga cga aca gcg gat gtc cgg gag acg act cgg cat gca ctc gac ctg Gly Arg Thr Ala Asp Val Arg Glu Thr Thr Arg His Ala Leu Asp Leu 1285 1290 1295	3888
gta caa ggc tgg ctt tcc gat cag cga ctc aac gaa tcc cga ttg ctc Val Gln Gly Trp Leu Ser Asp Gln Arg Leu Asn Glu Ser Arg Leu Leu 1300 1305 1310	3936
ctg gtg aca cag gga gca gtg gcc gtg gag ccg ggc gaa ccc gtg acc Leu Val Thr Gln Gly Ala Val Ala Val Glu Pro Gly Glu Pro Val Thr 1315 1320 1325	3984
gat ctg gcg cag gcc gcg ctc tgg gga ctg ctg cgg tcg acg cag acc Asp Leu Ala Gln Ala Ala Leu Trp Gly Leu Leu Arg Ser Thr Gln Thr 1330 1335 1340	4032
gaa cac cct gat cgc ttc gtc ctc gtc gat gtg cct gag ccc gcg caa Glu His Pro Asp Arg Phe Val Leu Val Asp Val Pro Glu Pro Ala Gln 1345 1350 1355 1360	4080
ctc ctc ccc gcg ctg ccg ggg gtg ctg gcc tgc ggc gaa cct cag ctc Leu Leu Pro Ala Leu Pro Gly Val Leu Ala Cys Gly Glu Pro Gln Leu 1365 1370 1375	4128
gcg ttg cga cgt ggc ggc gct cat gcg ccc aga ctg gct gga ctg ggc Ala Leu Arg Arg Gly Gly Ala His Ala Pro Arg Leu Ala Gly Leu Gly 1380 1385 1390	4176
agc gat gac gtc ctg ccc gtg ccg gac ggc acc ggg tgg cga ttg gag Ser Asp Asp Val Leu Pro Val Pro Asp Gly Thr Gly Trp Arg Leu Glu 1395 1400 1405	4224
gcc acg cgc ccg gga agc ctg gat ggg ttg gca ttg gtg gac gaa ccg Ala Thr Arg Pro Gly Ser Leu Asp Gly Leu Ala Leu Val Asp Glu Pro 1410 1415 1420	4272

acg gcc acg gca ccg ctg ggt gac ggt gag gtc agg att gcg atg cgc 4320  
 Thr Ala Thr Ala Pro Leu Gly Asp Gly Glu Val Arg Ile Ala Met Arg  
 1425 1430 1435 1440

gcg gcc ggg gtg aac ttc cgg gat gcg ctc atc gcg ctc ggt atg tat 4368  
 Ala Ala Gly Val Asn Phe Arg Asp Ala Leu Ile Ala Leu Gly Met Tyr  
 1445 1450 1455

ccc ggt gtg gca tcg ctg ggc agt gag ggc gcc ggg gtc gtg gtg gag 4416  
 Pro Gly Val Ala Ser Leu Gly Ser Glu Gly Ala Gly Val Val Val Glu  
 1460 1465 1470

acc ggc ccc ggc gtc acc ggc ctg gca ccc ggc gac cgc gtg atg gga 4464  
 Thr Gly Pro Gly Val Thr Gly Leu Ala Pro Gly Asp Arg Val Met Gly  
 1475 1480 1485

atg atc ccg aag gcg ttc ggg ccg ctc gcg gtc gcc gac cat cgc atg 4512  
 Met Ile Pro Lys Ala Phe Gly Pro Leu Ala Val Ala Asp His Arg Met  
 1490 1495 1500

gtg acg agg att ccc gct ggt tgg agc ttc gcg cgg gcc gca tcg gtg 4560  
 Val Thr Arg Ile Pro Ala Gly Trp Ser Phe Ala Arg Ala Ala Ser Val  
 1505 1510 1515 1520

ccg atc gtc ttt ctc acc gcc tac tac gcg ctg gtt gat ctc gcc ggg 4608  
 Pro Ile Val Phe Leu Thr Ala Tyr Tyr Ala Leu Val Asp Leu Ala Gly  
 1525 1530 1535

ttg aga cca ggg gag tcg ttg ctg gtt cat tcg gcc gcc ggt ggg gtg 4656  
 Leu Arg Pro Gly Glu Ser Leu Leu Val His Ser Ala Ala Gly Gly Val  
 1540 1545 1550

ggg atg gcc gcg atc caa ctc gcc agg cac ctc ggt gca gag gtg tac 4704  
 Gly Met Ala Ala Ile Gln Leu Ala Arg His Leu Gly Ala Glu Val Tyr  
 1555 1560 1565

gcc acc gct agc gag gac aag tgg caa gcc gtg gag ctg agc cga gaa 4752  
 Ala Thr Ala Ser Glu Asp Lys Trp Gln Ala Val Glu Leu Ser Arg Glu  
 1570 1575 1580

cac ctc gct tcg tcg cgg acg tgc gat ttc gag cag cag ttc ctc ggg 4800  
 His Leu Ala Ser Ser Arg Thr Cys Asp Phe Glu Gln Gln Phe Leu Gly  
 1585 1590 1595 1600

gca acc ggc gga cgc ggc gtc gac gtc gtg ctc aac tcc ctc gcc ggg 4848  
 Ala Thr Gly Gly Arg Gly Val Asp Val Val Leu Asn Ser Leu Ala Gly  
 1605 1610 1615

gag ttc gcc gat ggc tct ctg cga atg ctg ccg cgc ggc ggc cgt ttc	4896
Glu Phe Ala Asp Ala Ser Leu Arg Met Leu Pro Arg Gly Gly Arg Phe	
1620 1625 1630	
ctg gag ttg ggg aag acg gat gtt cgt gac ccc gtc gag gtc gcc gat	4944
Leu Glu Leu Gly Lys Thr Asp Val Arg Asp Pro Val Glu Val Ala Asp	
1635 1640 1645	
gcg cat ccg ggc gtg tct tac cag gct ttc gat acc gta gag gca ggc	4992
Ala His Pro Gly Val Ser Tyr Gln Ala Phe Asp Thr Val Glu Ala Gly	
1650 1655 1660	
ccg cag cga atc ggc gag atg ctt cac gag ctg gtg gag ttg ttc gag	5040
Pro Gln Arg Ile Gly Glu Met Leu His Glu Leu Val Glu Leu Phe Glu	
1665 1670 1675 1680	
gga cgc gtg ctg gag ccc ctg cct gtc acg gct tgg gac gtt cgg cag	5088
Gly Arg Val Leu Glu Pro Leu Pro Val Thr Ala Trp Asp Val Arg Gln	
1685 1690 1695	
gcg ccc gag gcg cta cgg cac ctg agc caa gcg cgg cat gtg gga aag	5136
Ala Pro Glu Ala Leu Arg His Leu Ser Gln Ala Arg His Val Gly Lys	
1700 1705 1710	
ctg gtg ctc acc atg cct ccg gtg tgg gac gcc gca ggc acg gtt ctg	5184
Leu Val Leu Thr Met Pro Pro Val Trp Asp Ala Ala Gly Thr Val Leu	
1715 1720 1725	
gtt acc ggc gga acg gga gca ctt ggc gca gag gtc gcc cgg cac ctc	5232
Val Thr Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg His Leu	
1730 1735 1740	
gtg atc gag cgc ggg gtg cga aac ctg gtc ctc gtc agc agg cgc ggt	5280
Val Ile Glu Arg Gly Val Arg Asn Leu Val Leu Val Ser Arg Arg Gly	
1745 1750 1755 1760	
ccc gca gcc agt ggc gct gct gag ctc gtg gcg caa ctg acg gcc tac	5328
Pro Ala Ala Ser Gly Ala Ala Glu Leu Val Ala Gln Leu Thr Ala Tyr	
1765 1770 1775	
ggt gcc gag gtt tcc ttg cag gct tgc gat gtc gcc gat cgt gag acc	5376
Gly Ala Glu Val Ser Leu Gln Ala Cys Asp Val Ala Asp Arg Glu Thr	
1780 1785 1790	
ttg gcg aag gtg ctt gcc agc atc ccg gac gag cat ccg ttg acc gcc	5424
Leu Ala Lys Val Leu Ala Ser Ile Pro Asp Glu His Pro Leu Thr Ala	
1795 1800 1805	

gtg gtg cac gcg gct ggt gtt ctc gac gac gga gtg tcc gaa tcg ctc	5472
Val Val His Ala Ala Gly Val Leu Asp Asp Gly Val Ser Glu Ser Leu	
1810 1815 1820	
acc gtg gag cgg ctg gac cag gtt ctg cgc ccg aag gtc gat ggc gcg	5520
Thr Val Glu Arg Leu Asp Gln Val Leu Arg Pro Lys Val Asp Gly Ala	
1825 1830 1835 1840	
cgg aat ctg ctc gag ctg atc gac ccg gac gtg gcc ctc gtg ttg ttc	5568
Arg Asn Leu Leu Glu Leu Ile Asp Pro Asp Val Ala Leu Val Leu Phe	
1845 1850 1855	
tcg tcg gtg tcg ggt gtg ctc ggc agc ggt ggg cag ggt aac tac gcg	5616
Ser Ser Val Ser Gly Val Leu Gly Ser Gly Gly Gln Gly Asn Tyr Ala	
1860 1865 1870	
gcg gcc aac tcc ttc ctc gac gca ttg gcg cag caa agg cag tcg cgc	5664
Ala Ala Asn Ser Phe Leu Asp Ala Leu Ala Gln Gln Arg Gln Ser Arg	
1875 1880 1885	
ggc cta ccg acg aga tca ttg gcc tgg ggg ccc tgg gcg gaa cat ggc	5712
Gly Leu Pro Thr Arg Ser Leu Ala Trp Gly Pro Trp Ala Glu His Gly	
1890 1895 1900	
atg gcc agc acc ttg cgc gaa gcc gag cag gat cga ttg gcg cga tct	5760
Met Ala Ser Thr Leu Arg Glu Ala Glu Gln Asp Arg Leu Ala Arg Ser	
1905 1910 1915 1920	
ggg ttg ctg ccg atc tcg acc gag gag ggg ttg tcc cag ttc gac gcc	5808
Gly Leu Leu Pro Ile Ser Thr Glu Glu Gly Leu Ser Gln Phe Asp Ala	
1925 1930 1935	
gcg tgc ggc ggc gcg cat acc gtg gtg gcg ccg gtt cga ttc agc cgc	5856
Ala Cys Gly Gly Ala His Thr Val Val Ala Pro Val Arg Phe Ser Arg	
1940 1945 1950	
ttg tcc gac ggg aac gcg atc aag ttc tcc gtc ctg caa ggt ttg gtc	5904
Leu Ser Asp Gly Asn Ala Ile Lys Phe Ser Val Leu Gln Gly Leu Val	
1955 1960 1965	
ggg ccg cat cgc gtc aac aaa gcg gcg act gcg gat gat gcc gag agc	5952
Gly Pro His Arg Val Asn Lys Ala Ala Thr Ala Asp Asp Ala Glu Ser	
1970 1975 1980	
ctc cgg aaa cgg ttg gga cgc ttg ccg gat gca gaa caa cat cgg att	6000
Leu Arg Lys Arg Leu Gly Arg Leu Pro Asp Ala Glu Gln His Arg Ile	
1985 1990 1995 2000	

ctg ctg gac ctc gcc cgc atg cat gtg gcg gca gtg ccc gga ttc gcc Leu Leu Asp Leu Val Arg Met His Val Ala Ala Val Leu Gly Phe Ala	6048
2005 2010 2015	
ggt tct cag gag atc acc gcg gac ggc acg ttc aag gtg ctg ggc ttc Gly Ser Gln Glu Ile Thr Ala Asp Gly Thr Phe Lys Val Leu Gly Phe	6096
2020 2025 2030	
gac tcg ttg acc gtg gtc gag ttg cgc aac cgg atc aac ggg gcg acg Asp Ser Leu Thr Val Val Glu Leu Arg Asn Arg Ile Asn Gly Ala Thr	6144
2035 2040 2045	
ggg ctg cga ctg ccc gcc acc ctg gtg ttc aac tac ccg acg ccg gat Gly Leu Arg Leu Pro Ala Thr Leu Val Phe Asn Tyr Pro Thr Pro Asp	6192
2050 2055 2060	
gcg ctc gcc gcg cac ctc gtc acc gcg ctg tcc gca gac cgc ctg gcc Ala Leu Ala Ala His Leu Val Thr Ala Leu Ser Ala Asp Arg Leu Ala	6240
2065 2070 2075 2080	
ggg aca ttc gag gaa ctc gac agg tgg gcg gcg aac ctg ccc acg ctg Gly Thr Phe Glu Glu Leu Asp Arg Trp Ala Ala Asn Leu Pro Thr Leu	6288
2085 2090 2095	
gcc agg gat gag gcc acg cgg gcg cag atc acc acc cgg cta cag gcg Ala Arg Asp Glu Ala Thr Arg Ala Gln Ile Thr Thr Arg Leu Gln Ala	6336
2100 2105 2110	
atc ttg cag agc ctg gcg gac gtg tcc ggc gga acc ggc ggc ggc tcc Ile Leu Gln Ser Leu Ala Asp Val Ser Gly Gly Thr Gly Gly Gly Ser	6384
2115 2120 2125	
gtg ccg gac cgg ctc aga tcg gcc acg gac gac gag ctt ttc caa ctc Val Pro Asp Arg Leu Arg Ser Ala Thr Asp Asp Glu Leu Phe Gln Leu	6432
2130 2135 2140	
ctc gac aac gat ctc gaa ctt ccc tga Leu Asp Asn Asp Leu Glu Leu Pro	6459
2145 2150	
<210> 44	
<211> 2152	
<212> PRT	
<213> Saccharopolyspora spinosa	
<400> 44	
Met Thr Val Thr Thr Ser Tyr Glu Glu Val Val Glu Ala Leu Arg Ala	

1	5	10	15
Ser Leu Lys Glu Asn Glu Arg Leu Arg Arg Gly Arg Asp Arg Phe Ser	20	25	30
Ala Glu Lys Asp Asp Pro Ile Ala Ile Val Ala Met Ser Cys Arg Tyr	35	40	45
Pro Gly Gln Val Ser Ser Pro Glu Asp Leu Trp Gln Leu Ala Ala Gly	50	55	60
Gly Val Asp Ala Ile Ser Glu Val Pro Gly Asp Arg Gly Trp Asp Leu	65	70	75
Asp Gly Val Phe Val Pro Asp Ser Asp Arg Pro Gly Thr Ser Tyr Ala	85	90	95
Cys Ala Gly Gly Phe Leu Gln Gly Val Ser Glu Phe Asp Ala Gly Phe	100	105	110
Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg	115	120	125
Leu Leu Leu Glu Val Ala Trp Glu Val Phe Glu Arg Ala Gly Leu Glu	130	135	140
Gln Arg Ser Thr Arg Gly Ser Arg Val Gly Val Phe Val Gly Thr Asn	145	150	155
Gly Gln Asp Tyr Ala Ser Trp Leu Arg Thr Pro Pro Pro Ala Val Ala	165	170	175
Gly His Val Leu Thr Gly Gly Ala Ala Ala Val Leu Ser Gly Arg Val	180	185	190
Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala	195	200	205
Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu Arg	210	215	220
Ala Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser	225	230	235
Thr Pro Lys Val Phe Leu Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro	245	250	255
Asp Gly Arg Cys Lys Ser Phe Ala Ala Gly Ala Asp Gly Thr Gly Trp			

260

265

270

Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg  
275 280 285

Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn  
290 295 300

Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Ser Ser Gln  
305 310 315 320

Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser  
325 330 335

Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp  
340 345 350

Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg Asp  
355 360 365

Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His  
370 375 380

Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala  
385 390 395 400

Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Glu Ser Pro Ser  
405 410 415

Pro Glu Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu Asn  
420 425 430

Thr Pro Trp Pro Arg Ser Gly Arg Val Arg Arg Val Gly Val Ser Ser  
435 440 445

Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Pro Pro  
450 455 460

Gly Val Pro Ser Gln Ser Ala Gly Pro Gly Ser Gly Ser Val Val Asp  
465 470 475 480

Val Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala Leu  
485 490 495

Ser Ala Gln Ala Thr Ala Leu Met Thr Tyr Leu Asp Glu Arg Pro Asp  
500 505 510

Val Ser Ser Leu Asp Val Gly Tyr Ser Leu Ala Leu Thr Arg Ser Ala



515

520

525

Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Glu Thr Leu Leu  
530 535 540

Cys Gly Val Lys Ala Leu Ser Ala Gly His Glu Ala Ser Gly Leu Val  
545 550 555 560

Thr Gly Ser Val Gly Ala Gly Gly Arg Ile Gly Phe Val Phe Ser Gly  
565 570 575

Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu Tyr Arg Ala Phe  
580 585 590

Pro Val Phe Ala Ala Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala  
595 600 605

His Leu Gly Gln Glu Ile Gly Val Arg Glu Val Val Ser Gly Ser Asp  
610 615 620

Ala Gln Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Phe Ala  
625 630 635 640

Leu Gln Val Gly Leu Leu Lys Leu Leu Asp Ser Trp Gly Val Arg Pro  
645 650 655

Ser Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala  
660 665 670

Ala Gly Val Val Ser Leu Ser Gly Ala Ala Arg Leu Val Ala Gly Arg  
675 680 685

Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Gly Met Leu Ala Val  
690 695 700

Pro Ala Gly Glu Glu Leu Leu Trp Ser Leu Leu Ala Asp Gln Gly Asp  
705 710 715 720

Arg Val Gly Ile Ala Ala Val Asn Ala Ala Gly Ser Val Val Leu Ser  
725 730 735

Gly Asp Arg Asp Val Leu Asp Asp Leu Ala Gly Arg Leu Asp Gly Gln  
740 745 750

Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser Tyr  
755 760 765

Arg Met Asp Pro Met Leu Ala Glu Phe Ala Glu Leu Ala Arg Thr Val

770

775

780

Asp Tyr Arg Arg Cys Glu Val Pro Ile Val Ser Thr Leu Thr Gly Asp  
785 790 795 800

Leu Asp Asp Ala Gly Arg Met Ser Gly Pro Asp Tyr Trp Val Arg Gln  
805 810 815

Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Gln Ala Leu Val Glu  
820 825 830

His Asp Val Ala Thr Val Val Glu Leu Gly Pro Asp Gly Ala Leu Ser  
835 840 845

Ala Leu Ile Gln Glu Cys Val Ala Ala Ser Asp His Ala Gly Arg Leu  
850 855 860

Ser Ala Val Pro Ala Met Arg Arg Asn Gln Asp Glu Ala Gln Lys Val  
865 870 875 880

Met Thr Ala Leu Ala His Val His Val Arg Gly Gly Ala Val Asp Trp  
885 890 895

Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Ile Glu Leu Pro Thr  
900 905 910

Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Val Pro Ser Asp Ser Gly  
915 920 925

Asp Val Thr Gly Ala Gly Leu Ala Gly Ala Glu His Pro Leu Leu Gly  
930 935 940

Ala Val Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg  
945 950 955 960

Ile Ser Val Arg Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly  
965 970 975

Glu Val Ile Val Ala Gly Thr Ala Leu Leu Glu Ile Ala Leu His Ala  
980 985 990

Gly Glu Arg Leu Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala  
995 1000 1005

Pro Leu Val Leu Pro Glu Arg Gly Ala Ile Gln Val Gln Leu Arg Val  
1010 1015 1020

Gly Ala Pro Glu Asn Ser Gly Arg Arg Pro Met Ala Leu Tyr Ser Arg

025	1030	1035	1040
Pro Glu Gly Ala Ala Glu His Asp Trp Thr Arg His Ala Thr Gly Arg	1045	1050	1055
Leu Ala Pro Gly Arg Gly Glu Ala Ala Gly Asp Leu Ala Asp Trp Pro	1060	1065	1070
Ala Pro Gly Ala Leu Pro Val Asp Leu Asp Glu Phe Tyr Arg Asp Leu	1075	1080	1085
Ala Glu Leu Gly Leu Glu Tyr Gly Pro Ile Phe Gln Gly Leu Lys Ala	1090	1095	1100
Ala Trp Arg Gln Gly Asp Glu Val Tyr Ala Glu Ala Ala Leu Pro Gly	1105	1110	1115
Thr Glu Asp Ser Gly Phe Gly Val His Pro Ala Leu Leu Asp Ala Ala	1125	1130	1135
Leu His Ala Thr Ala Val Arg Asp Met Asp Asp Ala Arg Leu Pro Phe	1140	1145	1150
Gln Trp Glu Gly Val Ser Leu His Ala Lys Ala Ala Pro Ala Leu Arg	1155	1160	1165
Val Arg Val Val Pro Ala Gly Asp Asp Ala Lys Ser Leu Leu Val Cys	1170	1175	1180
Asp Gly Thr Gly Arg Pro Val Ile Ser Val Asp Arg Leu Val Leu Arg	1185	1190	1195
Ser Ala Ala Ala Arg Arg Thr Gly Ala Arg Arg Gln Ala His Gln Ala	1205	1210	1215
Arg Leu Tyr Arg Leu Ser Trp Pro Thr Val Gln Leu Pro Thr Ser Ala	1220	1225	1230
Gln Pro Pro Ser Cys Val Leu Leu Gly Thr Ser Glu Val Ser Ala Asp	1235	1240	1245
Ile Gln Val Tyr Pro Asp Leu Arg Ser Leu Thr Ala Ala Leu Asp Ala	1250	1255	1260
Gly Ala Glu Pro Pro Gly Val Val Ile Ala Pro Thr Pro Pro Gly Gly	1265	1270	1275
Gly Arg Thr Ala Asp Val Arg Glu Thr Thr Arg His Ala Leu Asp Leu			

1285

1290

1295

Val Gln Gly Trp Leu Ser Asp Gln Arg Leu Asn Glu Ser Arg Leu Leu  
1300 1305 1310

Leu Val Thr Gln Gly Ala Val Ala Val Glu Pro Gly Glu Pro Val Thr  
1315 1320 1325

Asp Leu Ala Gln Ala Ala Leu Trp Gly Leu Leu Arg Ser Thr Gln Thr  
1330 1335 1340

Glu His Pro Asp Arg Phe Val Leu Val Asp Val Pro Glu Pro Ala Gln  
345 1350 1355 1360

Leu Leu Pro Ala Leu Pro Gly Val Leu Ala Cys Gly Glu Pro Gln Leu  
1365 1370 1375

Ala Leu Arg Arg Gly Gly Ala His Ala Pro Arg Leu Ala Gly Leu Gly  
1380 1385 1390

Ser Asp Asp Val Leu Pro Val Pro Asp Gly Thr Gly Trp Arg Leu Glu  
1395 1400 1405

Ala Thr Arg Pro Gly Ser Leu Asp Gly Leu Ala Leu Val Asp Glu Pro  
1410 1415 1420

Thr Ala Thr Ala Pro Leu Gly Asp Gly Glu Val Arg Ile Ala Met Arg  
425 1430 1435 1440

Ala Ala Gly Val Asn Phe Arg Asp Ala Leu Ile Ala Leu Gly Met Tyr  
1445 1450 1455

Pro Gly Val Ala Ser Leu Gly Ser Glu Gly Ala Gly Val Val Val Glu  
1460 1465 1470

Thr Gly Pro Gly Val Thr Gly Leu Ala Pro Gly Asp Arg Val Met Gly  
1475 1480 1485

Met Ile Pro Lys Ala Phe Gly Pro Leu Ala Val Ala Asp His Arg Met  
1490 1495 1500

Val Thr Arg Ile Pro Ala Gly Trp Ser Phe Ala Arg Ala Ala Ser Val  
505 1510 1515 1520

Pro Ile Val Phe Leu Thr Ala Tyr Tyr Ala Leu Val Asp Leu Ala Gly  
1525 1530 1535

Leu Arg Pro Gly Glu Ser Leu Leu Val His Ser Ala Ala Gly Gly Val

1540	1545	1550
Gly Met Ala Ala Ile Gln Leu Ala Arg His Leu Gly Ala Glu Val Tyr		
1555	1560	1565
Ala Thr Ala Ser Glu Asp Lys Trp Gln Ala Val Glu Leu Ser Arg Glu		
1570	1575	1580
His Leu Ala Ser Ser Arg Thr Cys Asp Phe Glu Gln Gln Phe Leu Gly		
585	1590	1595 1600
Ala Thr Gly Gly Arg Gly Val Asp Val Val Leu Asn Ser Leu Ala Gly		
1605	1610	1615
Glu Phe Ala Asp Ala Ser Leu Arg Met Leu Pro Arg Gly Gly Arg Phe		
1620	1625	1630
Leu Glu Leu Gly Lys Thr Asp Val Arg Asp Pro Val Glu Val Ala Asp		
1635	1640	1645
Ala His Pro Gly Val Ser Tyr Gln Ala Phe Asp Thr Val Glu Ala Gly		
1650	1655	1660
Pro Gln Arg Ile Gly Glu Met Leu His Glu Leu Val Glu Leu Phe Glu		
665	1670	1675 1680
Gly Arg Val Leu Glu Pro Leu Pro Val Thr Ala Trp Asp Val Arg Gln		
1685	1690	1695
Ala Pro Glu Ala Leu Arg His Leu Ser Gln Ala Arg His Val Gly Lys		
1700	1705	1710
Leu Val Leu Thr Met Pro Pro Val Trp Asp Ala Ala Gly Thr Val Leu		
1715	1720	1725
Val Thr Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg His Leu		
1730	1735	1740
Val Ile Glu Arg Gly Val Arg Asn Leu Val Leu Val Ser Arg Arg Gly		
745	1750	1755 1760
Pro Ala Ala Ser Gly Ala Ala Glu Leu Val Ala Gln Leu Thr Ala Tyr		
1765	1770	1775
Gly Ala Glu Val Ser Leu Gln Ala Cys Asp Val Ala Asp Arg Glu Thr		
1780	1785	1790
Leu Ala Lys Val Leu Ala Ser Ile Pro Asp Glu His Pro Leu Thr Ala		

1795

1800

1805

Val Val His Ala Ala Gly Val Leu Asp Asp Gly Val Ser Glu Ser Leu  
 1810 1815 1820

Thr Val Glu Arg Leu Asp Gln Val Leu Arg Pro Lys Val Asp Gly Ala  
 825 1830 1835 1840

Arg Asn Leu Leu Glu Leu Ile Asp Pro Asp Val Ala Leu Val Leu Phe  
 1845 1850 1855

Ser Ser Val Ser Gly Val Leu Gly Ser Gly Gly Gln Gly Asn Tyr Ala  
 1860 1865 1870

Ala Ala Asn Ser Phe Leu Asp Ala Leu Ala Gln Gln Arg Gln Ser Arg  
 1875 1880 1885

Gly Leu Pro Thr Arg Ser Leu Ala Trp Gly Pro Trp Ala Glu His Gly  
 1890 1895 1900

Met Ala Ser Thr Leu Arg Glu Ala Glu Gln Asp Arg Leu Ala Arg Ser  
 905 1910 1915 1920

Gly Leu Leu Pro Ile Ser Thr Glu Glu Gly Leu Ser Gln Phe Asp Ala  
 1925 1930 1935

Ala Cys Gly Gly Ala His Thr Val Val Ala Pro Val Arg Phe Ser Arg  
 1940 1945 1950

Leu Ser Asp Gly Asn Ala Ile Lys Phe Ser Val Leu Gln Gly Leu Val  
 1955 1960 1965

Gly Pro His Arg Val Asn Lys Ala Ala Thr Ala Asp Asp Ala Glu Ser  
 1970 1975 1980

Leu Arg Lys Arg Leu Gly Arg Leu Pro Asp Ala Glu Gln His Arg Ile  
 985 1990 1995 2000

Leu Leu Asp Leu Val Arg Met His Val Ala Ala Val Leu Gly Phe Ala  
 2005 2010 2015

Gly Ser Gln Glu Ile Thr Ala Asp Gly Thr Phe Lys Val Leu Gly Phe  
 2020 2025 2030

Asp Ser Leu Thr Val Val Glu Leu Arg Asn Arg Ile Asn Gly Ala Thr  
 2035 2040 2045

Gly Leu Arg Leu Pro Ala Thr Leu Val Phe Asn Tyr Pro Thr Pro Asp

2050

2055

2060

Ala Leu Ala Ala His Leu Val Thr Ala Leu Ser Ala Asp Arg Leu Ala  
 065 2070 2075 2080

Gly Thr Phe Glu Glu Leu Asp Arg Trp Ala Ala Asn Leu Pro Thr Leu  
 2085 2090 2095

Ala Arg Asp Glu Ala Thr Arg Ala Gln Ile Thr Thr Arg Leu Gln Ala  
 2100 2105 2110

Ile Leu Gln Ser Leu Ala Asp Val Ser Gly Gly Thr Gly Gly Gly Ser  
 2115 2120 2125

Val Pro Asp Arg Leu Arg Ser Ala Thr Asp Asp Glu Leu Phe Gln Leu  
 2130 2135 2140

Leu Asp Asn Asp Leu Glu Leu Pro  
 145 2150

&lt;210&gt; 45

&lt;211&gt; 9513

&lt;212&gt; DNA

&lt;213&gt; Saccharopolyspora spinosa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(9510)

&lt;223&gt; ORF20; Polyketidsynthase

&lt;400&gt; 45

atg tcg aat gaa gag aag ctc cgg gag tac ttg cgg cgt gcg ctc gtg 48  
 Met Ser Asn Glu Glu Lys Leu Arg Glu Tyr Leu Arg Arg Ala Leu Val  
 1 5 10 15

gat ctg cac cag gcg cgc gag cgg ctg cac gag gcg gag tcg gga gag 96  
 Asp Leu His Gln Ala Arg Glu Arg Leu His Glu Ala Glu Ser Gly Glu  
 20 25 30

cgg gaa ccc atc gcg atc gtg gcg atg ggc tgc cgg tac ccg ggt ggg 144  
 Arg Glu Pro Ile Ala Ile Val Ala Met Gly Cys Arg Tyr Pro Gly Gly  
 35 40 45

gtg cag gac ccg gaa ggg ctg tgg aaa ctg gtc gcc tcc ggt ggc gac 192  
 Val Gln Asp Pro Glu Gly Leu Trp Lys Leu Val Ala Ser Gly Gly Asp  
 50 55 60





tgt aag tcg ttc gcg gcg gct gcg gat gga acc ggt tgg ggt gag ggt	816
Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly	
260 265 270	
gcc ggt ttg gtg ttg ctg gag cgg ctg tca gac gcc agg cgg aac ggg	864
Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly	
275 280 285	
cat cgg gta ctg gcg gtt gtt cgt ggt agc gcg gtg aat cag gac ggt	912
His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly	
290 295 300	
gcg tcg aac gga ttg acg gcc ccg aac ggg ctg gcc cag gag cgg gtc	960
Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Leu Ala Gln Glu Arg Val	
305 310 315 320	
att cag cag gtg ctc acg agt gcg ggg ctg tcg gcg tcc gat gtg gac	1008
Ile Gln Gln Val Leu Thr Ser Ala Gly Leu Ser Ala Ser Asp Val Asp	
325 330 335	
gct gtg gag gcg cat gga acg ggt acg cgg ctt ggt gat ccg atc gag	1056
Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu	
340 345 350	
gcg cag gct ctg ata gcc gcc tat gga cag gat cgg gac cgg gac cgg	1104
Ala Gln Ala Leu Ile Ala Ala Tyr Gly Gln Asp Arg Asp Arg Asp Arg	
355 360 365	
ccg ctg tgg ttg ggg tcg gtc aag tcc aac atc ggt cat acg cag gcg	1152
Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala	
370 375 380	
gct gcg ggc gtc gct ggt gtg atc aag atg gtc atg gcg atg cgg cac	1200
Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala Met Arg His	
385 390 395 400	
ggg gag ctg ccg cgc acg ttg cac gtg gac gag ccg aat tcg cac gtg	1248
Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro Asn Ser His Val	
405 410 415	
gac tgg tcg gct ggt gcg gtc cga ctc ctg acc gag aac atc cgc tgg	1296
Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Glu Asn Ile Arg Trp	
420 425 430	
cca ggg acg ggt acg cgc cgc gct gga gtg tcg tcg ttc ggg gta agc	1344
Pro Gly Thr Gly Thr Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser	
435 440 445	

ggt acc aac gca cac gtc atc ctc gaa cac gac ccg ctc gcc gtg acc 1392  
 Gly Thr Asn Ala His Val Ile Leu Glu His Asp Pro Leu Ala Val Thr  
 450 455 460

gag aac gag gaa gca gcg cag tcc cca gca cct ggg atc gtg ccc tgg 1440  
 Glu Asn Glu Glu Ala Ala Gln Ser Pro Ala Pro Gly Ile Val Pro Trp  
 465 470 475 480

gcg ttg tcc ggg cgg tcg tcg acg gcg ctg cgg gcc cag gcc gaa cgg 1488  
 Ala Leu Ser Gly Arg Ser Ser Thr Ala Leu Arg Ala Gln Ala Glu Arg  
 485 490 495

ctg cgc gag ctg tgc gag cag acc gat ccc gac ccc gtc gat gtc ggt 1536  
 Leu Arg Glu Leu Cys Glu Gln Thr Asp Pro Asp Pro Val Asp Val Gly  
 500 505 510

ctc tca ctg gcc gcc acg cgc acg gct tgg gag cac cga gcg gtg gtg 1584  
 Phe Ser Leu Ala Ala Thr Arg Thr Ala Trp Glu His Arg Ala Val Val  
 515 520 525

ctt ggt cgg gac agc gct acg ttg cgc tcc ggg ctt ggc gtt gtt gcc 1632  
 Leu Gly Arg Asp Ser Ala Thr Leu Arg Ser Gly Leu Gly Val Val Ala  
 530 535 540

agc ggt gaa cca gcg gtc gat gtc gtt gag ggg agc gtc ctg gac ggc 1680  
 Ser Gly Glu Pro Ala Val Asp Val Val Glu Gly Ser Val Leu Asp Gly  
 545 550 555 560

gag gtc gtc ttc gtc ttc ccc ggt cag ggc tgg cag tgg gcc ggt atg 1728  
 Glu Val Val Phe Val Phe Pro Gly Gln Gly Trp Gln Trp Ala Gly Met  
 565 570 575

gca gtc gac ctg ctg gac gct tcg ccg acg ttc gcg cgc cac atg gac 1776  
 Ala Val Asp Leu Leu Asp Ala Ser Pro Thr Phe Ala Arg His Met Asp  
 580 585 590

gag tgc gcc acc gcg ctg cgg agg tac gtg gac tgg tcg ttg gtc gac 1824  
 Glu Cys Ala Thr Ala Leu Arg Arg Tyr Val Asp Trp Ser Leu Val Asp  
 595 600 605

gtg ctg cgc gga gcg gag aac tcc cca ccg ctg gac cgg gtg gac gtg 1872  
 Val Leu Arg Gly Ala Glu Asn Ser Pro Pro Leu Asp Arg Val Asp Val  
 610 615 620

ctc cag ccc gcg tcc ttc gcg gtg atg gtg tcg ctc gcc gag gtg tgg 1920  
 Leu Gln Pro Ala Ser Phe Ala Val Met Val Ser Leu Ala Glu Val Trp  
 625 630 635 640

cgt tcc tac ggg gtg agg ccg gcg gcc gtc gtc ggc cac agt caa ggc	1968
Arg Ser Tyr Gly Val Arg Pro Ala Ala Val Val Gly His Ser Gln Gly	
645 650 655	
gaa atc gcc gcg gcc tgc gca gcc ggg gtg ctg ccg ctg gag gat gcg	2016
Glu Ile Ala Ala Ala Cys Ala Ala Gly Val Leu Pro Leu Glu Asp Ala	
660 665 670	
gcc agg ctt gtc gca ttg cgc agc aga gcg ttg aag gga ctt tcg ggg	2064
Ala Arg Leu Val Ala Leu Arg Ser Arg Ala Leu Lys Gly Leu Ser Gly	
675 680 685	
cgg ggt ggc atg gcg tcg ctg gcc tgc cct gcg gat gag gtc gcg gca	2112
Arg Gly Gly Met Ala Ser Leu Ala Cys Pro Ala Asp Glu Val Ala Ala	
690 695 700	
ttg ttc gcg gga tcg ggc ggc cgt ctg gaa gtt gcg gcg atc aac ggc	2160
Leu Phe Ala Gly Ser Gly Gly Arg Leu Glu Val Ala Ala Ile Asn Gly	
705 710 715 720	
ccg cga tcg gtc gtg gtg tcc ggc gat ctg gaa gcg gtg gac gaa ctg	2208
Pro Arg Ser Val Val Val Ser Gly Asp Leu Glu Ala Val Asp Glu Leu	
725 730 735	
ctg gca gag tgc gct gaa aag gac atg cgt gca cgc cgt atc ccc gtc	2256
Leu Ala Glu Cys Ala Glu Lys Asp Met Arg Ala Arg Arg Ile Pro Val	
740 745 750	
gac tac gcc tcg cat tca gcg cac gtg gag gtg gtt cgg agc ccg gtg	2304
Asp Tyr Ala Ser His Ser Ala His Val Glu Val Val Arg Ser Pro Val	
755 760 765	
ctg gcg gcc gcc gcc ggg gtg cga cac ccg gac ggc cag gtg ccg tgg	2352
Leu Ala Ala Ala Ala Gly Val Arg His Arg Asp Gly Gln Val Pro Trp	
770 775 780	
tgg tcg acg gtg atc ggc gac tgg gtg gat ccg gcc agg ctg gac ggc	2400
Trp Ser Thr Val Ile Gly Asp Trp Val Asp Pro Ala Arg Leu Asp Gly	
785 790 795 800	
gag tat tgg tat cgg aac ctc ccg cag ccg gtc ccg ttc gaa cac gcc	2448
Glu Tyr Trp Tyr Arg Asn Leu Arg Gln Pro Val Arg Phe Glu His Ala	
805 810 815	
gtg cag ggc ctg gtc gag ccg gga ttc ggc ctg ttc atc gaa atg agt	2496
Val Gln Gly Leu Val Glu Arg Gly Phe Gly Leu Phe Ile Glu Met Ser	
820 825 830	

gcg cat ccg gtg ctg acc acg gcg gtc gag gaa acc ggt gcg gag tcg	2544
Ala His Pro Val Leu Thr Thr Ala Val Glu Glu Thr Gly Ala Glu Ser	
835 840 845	
gag acc gcc gtg gcc gcg gta ggt acc ttg cga cgt gac tcg ggc ggc	2592
Glu Thr Ala Val Ala Ala Val Gly Thr Leu Arg Arg Asp Ser Gly Gly	
850 855 860	
ctc cgg agg ttg ttg cat tcg ctg gcc gag gcg tac gtg cgc ggc gcc	2640
Leu Arg Arg Leu Leu His Ser Leu Ala Glu Ala Tyr Val Arg Gly Ala	
865 870 875 880	
acc gtg gac tgg gcc gtg gcg ttc ggg ggc gcg ggc cga cgg ctg gac	2688
Thr Val Asp Trp Ala Val Ala Phe Gly Gly Ala Gly Arg Arg Leu Asp	
885 890 895	
ctg ccg acc tac ccg ttc cag cgc cag cgg tac tgg ctg gac aag gga	2736
Leu Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Asp Lys Gly	
900 905 910	
gct gcc tcc gac gag gct cgt gcg gtc tcg gac ccg gcg gcg ggc tgg	2784
Ala Ala Ser Asp Glu Ala Arg Ala Val Ser Asp Pro Ala Ala Gly Trp	
915 920 925	
ttc tgg caa gcc gtg gcg cgc caa gac ctg aaa agc gtg tcc gat gcc	2832
Phe Trp Gln Ala Val Ala Arg Gln Asp Leu Lys Ser Val Ser Asp Ala	
930 935 940	
ctc gat ctc gac gcc gac gca ccg ctg agc gca aca ctt cca gcc ctg	2880
Leu Asp Leu Asp Ala Asp Ala Pro Leu Ser Ala Thr Leu Pro Ala Leu	
945 950 955 960	
tcc gtc tgg cac cgt cag gaa cga gaa agg gtc ttg gca gac ggt tgg	2928
Ser Val Trp His Arg Gln Glu Arg Glu Arg Val Leu Ala Asp Gly Trp	
965 970 975	
cgg tac cga gtc gac tgg gta cgg gtg gcc ccg cag ccg gtc cgg aga	2976
Arg Tyr Arg Val Asp Trp Val Arg Val Ala Pro Gln Pro Val Arg Arg	
980 985 990	
acg cgg gaa acc tgg ctc ctg gtc gtt ccc ccg ggc ggc atc gag gaa	3024
Thr Arg Glu Thr Trp Leu Leu Val Val Pro Pro Gly Gly Ile Glu Glu	
995 1000 1005	
gcg ctg gtc gaa cgg ctg acg gat gcg ttg aac acg cga ggg atc agc	3072
Ala Leu Val Glu Arg Leu Thr Asp Ala Leu Asn Thr Arg Gly Ile Ser	
1010 1015 1020	

acc ctg cgc ctc gac gtg cca ccg gcg gcg acc agt ggc gaa ctc gca	3120
Thr Leu Arg Leu Asp Val Pro Pro Ala Ala Thr Ser Gly Glu Leu Ala	
1025                      1030                      1035                      1040	
acc gaa ctc cgc gcc gca gcc gac ggt gac ccg gtg aag gca atc ctg	3168
Thr Glu Leu Arg Ala Ala Ala Asp Gly Asp Pro Val Lys Ala Ile Leu	
1045                      1050                      1055	
tcg ctc acc gcg ttg gac gag cga ccc cac ccc gaa tgc aag gac gtc	3216
Ser Leu Thr Ala Leu Asp Glu Arg Pro His Pro Glu Cys Lys Asp Val	
1060                      1065                      1070	
ccg agc ggg att gcc ttg ctg ctg aac ctg gtc aag gcg ctc ggt gaa	3264
Pro Ser Gly Ile Ala Leu Leu Leu Asn Leu Val Lys Ala Leu Gly Glu	
1075                      1080                      1085	
gcc gac ctc aga att cct ctg tgg acc atc acg cgt ggt gcg gtc aag	3312
Ala Asp Leu Arg Ile Pro Leu Trp Thr Ile Thr Arg Gly Ala Val Lys	
1090                      1095                      1100	
gca ggc ccc gca gat cgg ctg ctg cgc ccg atg cag gcg caa gca tgg	3360
Ala Gly Pro Ala Asp Arg Leu Leu Arg Pro Met Gln Ala Gln Ala Trp	
1105                      1110                      1115                      1120	
ggt ctg ggg cga gta gcc gca ctc gaa cac ccc gag cgc tgg ggt ggg	3408
Gly Leu Gly Arg Val Ala Ala Leu Glu His Pro Glu Arg Trp Gly Gly	
1125                      1130                      1135	
ctg atc gac ctg ccg gat tcg ctg gac ggc gac gtc ctc acg agg ctg	3456
Leu Ile Asp Leu Pro Asp Ser Leu Asp Gly Asp Val Leu Thr Arg Leu	
1140                      1145                      1150	
ggc gaa gcg ctc acc aac ggc ttg gcg gaa gac caa ctg gcg att cgc	3504
Gly Glu Ala Leu Thr Asn Gly Leu Ala Glu Asp Gln Leu Ala Ile Arg	
1155                      1160                      1165	
cag tcg ggc gtg ctg gcc ccg cga ctg gta ccc gcc ccg gcg aat cag	3552
Gln Ser Gly Val Leu Ala Arg Arg Leu Val Pro Ala Pro Ala Asn Gln	
1170                      1175                      1180	
ccc gct gga cgt aag tgg cgc ccc cga ggg agc gcg ctg atc acg ggc	3600
Pro Ala Gly Arg Lys Trp Arg Pro Arg Gly Ser Ala Leu Ile Thr Gly	
1185                      1190                      1195                      1200	
gga ctc ggc gcg gtg ggc gca cag gtg gcg agg tgg ttg gcc gaa atc	3648
Gly Leu Gly Ala Val Gly Ala Gln Val Ala Arg Trp Leu Ala Glu Ile	
1205                      1210                      1215	

gga gcc gag cga atc gtg ctc acc agt cga cgg ggc aac caa gca gca	3696
Gly Ala Glu Arg Ile Val Leu Thr Ser Arg Arg Gly Asn Gln Ala Ala	
1220 1225 1230	
ggc gcc gcc gag ctg gaa gcc gaa ctc cgg gcc ctt gga gcg caa gtg	3744
Gly Ala Ala Glu Leu Glu Ala Glu Leu Arg Ala Leu Gly Ala Gln Val	
1235 1240 1245	
tcc atc gtg gct tgc gac gtg acc gat cgt gcc gag atg tcc gca cta	3792
Ser Ile Val Ala Cys Asp Val Thr Asp Arg Ala Glu Met Ser Ala Leu	
1250 1255 1260	
ctg gcc gag ttc gac gtc acc gcg gtg ttc cac gcg gcc gga gtc ggt	3840
Leu Ala Glu Phe Asp Val Thr Ala Val Phe His Ala Ala Gly Val Gly	
1265 1270 1275 1280	
ggc ctg ctg ccg ttg gcg gag acc gac cag aac ggc ctg gcc gaa ata	3888
Arg Leu Leu Pro Leu Ala Glu Thr Asp Gln Asn Gly Leu Ala Glu Ile	
1285 1290 1295	
tgc gcg gcg aag gtc cgc gcc gct cag gtg ctg gac gaa ctg tgc gac	3936
Cys Ala Ala Lys Val Arg Gly Ala Gln Val Leu Asp Glu Leu Cys Asp	
1300 1305 1310	
agc acc gat ctc gat gcc ttc gtc ctg ttc tcc tcg ggt gcc ggg gta	3984
Ser Thr Asp Leu Asp Ala Phe Val Leu Phe Ser Ser Gly Ala Gly Val	
1315 1320 1325	
tgg ggc ggg gcc ggt cag gcc gct tac gcc gcg gcg aac gca ttc ttg	4032
Trp Gly Gly Gly Gly Gln Gly Ala Tyr Gly Ala Ala Asn Ala Phe Leu	
1330 1335 1340	
gac aca ctc gcc gaa caa cgc cga gca cgc ggt ctg ccg gca acc tcg	4080
Asp Thr Leu Ala Glu Gln Arg Arg Ala Arg Gly Leu Pro Ala Thr Ser	
1345 1350 1355 1360	
atc tcc tgg gcc agt tgg gcc gcc gcc gcc atg gcc gac gcc gcg gcg	4128
Ile Ser Trp Gly Ser Trp Ala Gly Gly Gly Met Ala Asp Gly Ala Ala	
1365 1370 1375	
ggc gaa cac ctg cgg cga cgc ggg ata cgt ccg atg ccg gcg gcg tcg	4176
Gly Glu His Leu Arg Arg Arg Gly Ile Arg Pro Met Pro Ala Ala Ser	
1380 1385 1390	
gcc atc ctg gct ctg cag gaa gta ctt gac cag gat gag acg tgc gtg	4224
Ala Ile Leu Ala Leu Gln Glu Val Leu Asp Gln Asp Glu Thr Cys Val	
1395 1400 1405	

tcg atc gct gat gtg gac tgg gac cga ttc gtt ccc acg ttc gcc gcg	4272
Ser Ile Ala Asp Val Asp Trp Asp Arg Phe Val Pro Thr Phe Ala Ala	
1410 1415 1420	
act cgc gcc acc cgg ttg ttc gac gaa gtg ccg gcg gcg aga aag gcg	4320
Thr Arg Ala Thr Arg Leu Phe Asp Glu Val Pro Ala Ala Arg Lys Ala	
1425 1430 1435 1440	
atg ccc gcg aat ggg ccg gca gaa cca ggc ggc tcg ccg ttc gcc cgc	4368
Met Pro Ala Asn Gly Pro Ala Glu Pro Gly Gly Ser Pro Phe Ala Arg	
1445 1450 1455	
aat ctc gcg gag ctg ccg gaa gcc caa cga cgc cac gaa ctg gtg gat	4416
Asn Leu Ala Glu Leu Pro Glu Ala Gln Arg Arg His Glu Leu Val Asp	
1460 1465 1470	
ctg gtg tgc gcc cag gtg gca acc gtg ctc ggg cac ggc agt cgc gag	4464
Leu Val Cys Ala Gln Val Ala Thr Val Leu Gly His Gly Ser Arg Glu	
1475 1480 1485	
gaa gtc cag ccc gag cgg gcg ttc cgc gcg ctc ggg ttc gac tcc ctc	4512
Glu Val Gln Pro Glu Arg Ala Phe Arg Ala Leu Gly Phe Asp Ser Leu	
1490 1495 1500	
atg gcg gtg gat ctg cgc aat cgt ttg acc acc gcc acc ggg ttg cgc	4560
Met Ala Val Asp Leu Arg Asn Arg Leu Thr Thr Ala Thr Gly Leu Arg	
1505 1510 1515 1520	
ctg ccg acc aca acc gtc ttc gac tac ccg aat ccg gcc gcc ttg gcc	4608
Leu Pro Thr Thr Thr Val Phe Asp Tyr Pro Asn Pro Ala Ala Leu Ala	
1525 1530 1535	
gct cac ctg ctc gag gag ctg gtg ggt gat gtc gcg tcg gct gcg gtg	4656
Ala His Leu Leu Glu Glu Leu Val Gly Asp Val Ala Ser Ala Ala Val	
1540 1545 1550	
acc gct gcc agc gcg ccc gcg agt gac gaa ccg atc gcg atc gtc gcg	4704
Thr Ala Ala Ser Ala Pro Ala Ser Asp Glu Pro Ile Ala Ile Val Ala	
1555 1560 1565	
atg agc tgc cgg ttt ccg ggt ggc gcg cac tcg ccg gaa gac ctg tgg	4752
Met Ser Cys Arg Phe Pro Gly Gly Ala His Ser Pro Glu Asp Leu Trp	
1570 1575 1580	
cgg ctg gtc gcc gcc ggc acg gag gtg atc ggc gag ttc ccc tcc gac	4800
Arg Leu Val Ala Ala Gly Thr Glu Val Ile Gly Glu Phe Pro Ser Asp	
1585 1590 1595 1600	

cgg ggc tgg gat gcg gaa ggc ctt tac gat ccg gat gct tcc agg cct	4848
Arg Gly Trp Asp Ala Glu Gly Leu Tyr Asp Pro Asp Ala Ser Arg Pro	
1605 1610 1615	
gga acg acg tat gcg cgg atg gcg gga ttc ctc tac gac gcc ggt gag	4896
Gly Thr Thr Tyr Ala Arg Met Ala Gly Phe Leu Tyr Asp Ala Gly Glu	
1620 1625 1630	
ttc gat gcc gac ctg ttc ggc atc agc cca cgt gag gcg ttg gcg atg	4944
Phe Asp Ala Asp Leu Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met	
1635 1640 1645	
gat ccg cag cag cgg ttg gtg ctc gaa atc gcc tgg gaa gcc ctc gaa	4992
Asp Pro Gln Gln Arg Leu Val Leu Glu Ile Ala Trp Glu Ala Leu Glu	
1650 1655 1660	
cgg gcc gga atc gat ccg ttg tcc ttg aag ggc agt ggg gtc ggc acg	5040
Arg Ala Gly Ile Asp Pro Leu Ser Leu Lys Gly Ser Gly Val Gly Thr	
1665 1670 1675 1680	
tac atc ggc gct gga agc cgt ggg tac gcg acg gat gtg cgg cag ttt	5088
Tyr Ile Gly Ala Gly Ser Arg Gly Tyr Ala Thr Asp Val Arg Gln Phe	
1685 1690 1695	
ccc gag gag gcg gag ggc tac ctg ctg acg ggt acc tcg gcc agt gtg	5136
Pro Glu Glu Ala Glu Gly Tyr Leu Leu Thr Gly Thr Ser Ala Ser Val	
1700 1705 1710	
ctg tcg ggt cgg gtc gcg tat tcg ttt ggt ttc gag ggt cct gcg gtg	5184
Leu Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val	
1715 1720 1725	
acg gtg gat acg gct tgt tcg tcg tcg ttg gtg gcg ttg cat ctg gcg	5232
Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala	
1730 1735 1740	
tgc cag tcg ttg cgt tcg ggc gag tgt gat ctg gcg ttg gcc ggt ggt	5280
Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly	
1745 1750 1755 1760	
gtg acc gtg atg tcg acg ccg gag atg ttc gtg gag ttc tcc cgt cag	5328
Val Thr Val Met Ser Thr Pro Glu Met Phe Val Glu Phe Ser Arg Gln	
1765 1770 1775	
cgc ggt ttg gcg ccg gat ggg cgg tgc aag tcg ttc gcg gag agc gcg	5376
Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Glu Ser Ala	
1780 1785 1790	



gac ggc acc ggc tgg ggc gaa ggc gcg ggc ctg ttg ttg ctg gag cgg	5424
Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg	
1795 1800 1805	
ttg tcg gac gcc cac cgg aat ggg cat cgg gtg ttg gcg gtg gtt cgt	5472
Leu Ser Asp Ala His Arg Asn Gly His Arg Val Leu Ala Val Val Arg	
1810 1815 1820	
ggg tca gcg gtg aat cag gac ggc gcc tcg aac gga ctg gcg gcg ccg	5520
Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Ala Ala Pro	
1825 1830 1835 1840	
aac ggt ccg tcg cag cag cgg gtg atc aac cag gca ctc gcg aat gcg	5568
Asn Gly Pro Ser Gln Gln Arg Val Ile Asn Gln Ala Leu Ala Asn Ala	
1845 1850 1855	
gct ctt tcg gcg tcc gat gtg gat gcg gtg gag gca cat ggc acc ggg	5616
Ala Leu Ser Ala Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly	
1860 1865 1870	
acc agg ctg ggt gat ccg atc gag gcg cag gca ttg atc gca acg tat	5664
Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr	
1875 1880 1885	
ggg cag gcc cgg gag cgg gat cgg ccc ttg tgg ctg ggg tcg gtc aag	5712
Gly Gln Ala Arg Glu Arg Asp Arg Pro Leu Trp Leu Gly Ser Val Lys	
1890 1895 1900	
tcg aac atc ggt cat acg cag gcc gcg gcg ggt gtt gcc ggt gtg atc	5760
Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile	
1905 1910 1915 1920	
aag atg gtg atg gcc atg cgg cac ggg cag ctg ccc gcc tcg ctg cac	5808
Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro Ala Ser Leu His	
1925 1930 1935	
gcg gat gag ccc acg tcg gag gtc gat tgg tcg tcg ggg gcg gtc cgg	5856
Ala Asp Glu Pro Thr Ser Glu Val Asp Trp Ser Ser Gly Ala Val Arg	
1940 1945 1950	
ctc ctc gcc gaa cag gta cct tgg ccg gag tct gac cgt gtt cgt cgg	5904
Leu Leu Ala Glu Gln Val Pro Trp Pro Glu Ser Asp Arg Val Arg Arg	
1955 1960 1965	
gtg ggg gtt tcg tcg ttc ggg atc agc ggc acc aac gca cat gtg atc	5952
Val Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile	
1970 1975 1980	



ggt tcg tgg ggc gtt cgg ccg gat gtg gtg atg ggg cat tcg gtc ggg 6576  
 Gly Ser Trp Gly Val Arg Pro Asp Val Val Met Gly His Ser Val Gly  
 2180 2185 2190

gag ttg gcc gcc gcg ttt gcg gct ggc gtg ttg tcg ttg cgg gat gcg 6624  
 Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Leu Ser Leu Arg Asp Ala  
 2195 2200 2205

gct cgg ttg gtg gcc gcg cgc gcc cgg ttg atg caa gcc ctg ccc tct 6672  
 Ala Arg Leu Val Ala Ala Arg Ala Arg Leu Met Gln Ala Leu Pro Ser  
 2210 2215 2220

gac ggc gcg atg ttg gcg gtg gct gct ggt gaa gac ctt gtt cgg cca 6720  
 Asp Gly Ala Met Leu Ala Val Ala Ala Gly Glu Asp Leu Val Arg Pro  
 2225 2230 2235 2240

ttg ctg gcc ggt cgg gag gag tcc gtg agc gtc gcc gcg ctc aat gcc 6768  
 Leu Leu Ala Gly Arg Glu Glu Ser Val Ser Val Ala Ala Leu Asn Ala  
 2245 2250 2255

ccc ggt tcg gtg gtg ttg tcg ggc gat cgg gag gtg ctg gcc agc atc 6816  
 Pro Gly Ser Val Val Leu Ser Gly Asp Arg Glu Val Leu Ala Ser Ile  
 2260 2265 2270

gtc ggc cgg ctg acc gag ctc cga gtc cgg acg cgg cgc ttg cgg gtc 6864  
 Val Gly Arg Leu Thr Glu Leu Arg Val Arg Thr Arg Arg Leu Arg Val  
 2275 2280 2285

tcc cat gct ttt cat tcg cac cgg atg gac ccg atg ttg ggc gag ttc 6912  
 Ser His Ala Phe His Ser His Arg Met Asp Pro Met Leu Gly Glu Phe  
 2290 2295 2300

gcc cag atc gcc gag tct gcg gag ttc ggt aag cca acg aca ccg ctt 6960  
 Ala Gln Ile Ala Glu Ser Ala Glu Phe Gly Lys Pro Thr Thr Pro Leu  
 2305 2310 2315 2320

gtg tcg acg ttg acg ggt gag ctc gac aga gcc gcg gaa atg agc aca 7008  
 Val Ser Thr Leu Thr Gly Glu Leu Asp Arg Ala Ala Glu Met Ser Thr  
 2325 2330 2335

cca ggg tat tgg gtg cgc cag gcg cgt gaa ccc gtc cgt ttc gcc gac 7056  
 Pro Gly Tyr Trp Val Arg Gln Ala Arg Glu Pro Val Arg Phe Ala Asp  
 2340 2345 2350

ggt gtc cag gcc ctg gca gcg cag ggc ata ggc acg gtc gtc gag ctc 7104  
 Gly Val Gln Ala Leu Ala Ala Gln Gly Ile Gly Thr Val Val Glu Leu  
 2355 2360 2365

ggc ccg gac gga acg ctg gcg gca ctg gtt cgg gag tgt gcg acc gag	7152
Gly Pro Asp Gly Thr Leu Ala Ala Leu Val Arg Glu Cys Ala Thr Glu	
2370 2375 2380	
tcc gat cgg gtt ggg cgg att tcg tcg atc cca ctg atg cgc agg gag	7200
Ser Asp Arg Val Gly Arg Ile Ser Ser Ile Pro Leu Met Arg Arg Glu	
2385 2390 2395 2400	
cgg gac gag acc cgt tcg gtg atg aca gcc ctg gcg cat ctc cac acc	7248
Arg Asp Glu Thr Arg Ser Val Met Thr Ala Leu Ala His Leu His Thr	
2405 2410 2415	
cgt ggt ggt gag gtg gac tgg cag gcg ttt ttc gcc ggt acc ggc gct	7296
Arg Gly Gly Glu Val Asp Trp Gln Ala Phe Phe Ala Gly Thr Gly Ala	
2420 2425 2430	
agg cag ctc gag ttg cca acg tat gcc ttc caa cga cag cac tac tgg	7344
Arg Gln Leu Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln His Tyr Trp	
2435 2440 2445	
atc gag tcc agt gcg cgg cca gca cgc gac cgc gca gac atc ggc gag	7392
Ile Glu Ser Ser Ala Arg Pro Ala Arg Asp Arg Ala Asp Ile Gly Glu	
2450 2455 2460	
gtg gcg gaa cag ttc tgg acc gcg gtt gac caa ggc gat ctg gca acg	7440
Val Ala Glu Gln Phe Trp Thr Ala Val Asp Gln Gly Asp Leu Ala Thr	
2465 2470 2475 2480	
ttg gtc gcc gct ctg gat ctt ggg gcg gac gac gac aca tgc gca tcg	7488
Leu Val Ala Ala Leu Asp Leu Gly Ala Asp Asp Asp Thr Cys Ala Ser	
2485 2490 2495	
ctg agc gat gta ttg ccg gcg ttg tcc tcc tgg cga agc gga ctc cgc	7536
Leu Ser Asp Val Leu Pro Ala Leu Ser Ser Trp Arg Ser Gly Leu Arg	
2500 2505 2510	
aac cgt tcg ctc gtc gat tcc tgc cgg tac cga atc agt tgg cat tcc	7584
Asn Arg Ser Leu Val Asp Ser Cys Arg Tyr Arg Ile Ser Trp His Ser	
2515 2520 2525	
tct cgg gag gtg ccg gcc ccg aag att tcc ggt acc tgg ctg ttg gtc	7632
Ser Arg Glu Val Pro Ala Pro Lys Ile Ser Gly Thr Trp Leu Leu Val	
2530 2535 2540	
gtg ccc ggt gct gcg gat gac gga ttg gtc acg gct ttg acg agt tca	7680
Val Pro Gly Ala Ala Asp Asp Gly Leu Val Thr Ala Leu Thr Ser Ser	
2545 2550 2555 2560	

ctg gtc gga ggc ggc gcc gag gtc gtc cgg atc ggc ctg tcc gaa gag	7728
Leu Val Gly Gly Gly Ala Glu Val Val Arg Ile Gly Leu Ser Glu Glu	
2565 2570 2575	
gac ccg cac cgc gag gac gtc gca cag cgg ctg gcc aat gcg ctg acg	7776
Asp Pro His Arg Glu Asp Val Ala Gln Arg Leu Ala Asn Ala Leu Thr	
2580 2585 2590	
gat gcc ggt caa ctc ggt ggc gtg ctt tcg ctg ttg ggg ctc gat gaa	7824
Asp Ala Gly Gln Leu Gly Gly Val Leu Ser Leu Leu Gly Leu Asp Glu	
2595 2600 2605	
tcg cct gct ccg gga ttc tcc tgc ttg cca act ggt ttc gcg ctg act	7872
Ser Pro Ala Pro Gly Phe Ser Cys Leu Pro Thr Gly Phe Ala Leu Thr	
2610 2615 2620	
gtg cag ctt ctg cgg gcc ttg cgg aag gcc gac gtc gag gcg cct ttt	7920
Val Gln Leu Leu Arg Ala Leu Arg Lys Ala Asp Val Glu Ala Pro Phe	
2625 2630 2635 2640	
tgg gcg gtg acg cgc ggc ggc gtc gcg ttg gaa gat gta cgc gtg tct	7968
Trp Ala Val Thr Arg Gly Gly Val Ala Leu Glu Asp Val Arg Val Ser	
2645 2650 2655	
ccg gag cag gcc ctg gtc tgg ggg ctg ctg cgt gtc gcg gga ctg gag	8016
Pro Glu Gln Ala Leu Val Trp Gly Leu Leu Arg Val Ala Gly Leu Glu	
2660 2665 2670	
cac ccg gag ttc tgg ggt ggc ttg atc gac ctg cca tcg gac tgg gac	8064
His Pro Glu Phe Trp Gly Gly Leu Ile Asp Leu Pro Ser Asp Trp Asp	
2675 2680 2685	
gac cga ttg ggt gcc cgg ttg gcg ggt gtg ttg gcg gat ggt ggc gag	8112
Asp Arg Leu Gly Ala Arg Leu Ala Gly Val Leu Ala Asp Gly Gly Glu	
2690 2695 2700	
gat caa gtc gcc att cgc cgt ggt ggt gtg ttc gtg cgg cgg ttg gaa	8160
Asp Gln Val Ala Ile Arg Arg Gly Gly Val Phe Val Arg Arg Leu Glu	
2705 2710 2715 2720	
cgc gct ggt gcg tcg ggt gcc ggg tcg gtg tgg cgt cct cgg ggg acg	8208
Arg Ala Gly Ala Ser Gly Ala Gly Ser Val Trp Arg Pro Arg Gly Thr	
2725 2730 2735	
gtg ttg gtg acg ggt ggt acg ggc ggt ttg ggg gcg cat gtt gcc cgg	8256
Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala Arg	
2740 2745 2750	

tgg ttg gcc ggt gcc ggg gct gag cac gtg gtg ttg acc agc cgt cga	8304
Trp Leu Ala Gly Ala Gly Ala Glu His Val Val Leu Thr Ser Arg Arg	
2755 2760 2765	
gga gcg gac gct ccg ggc gct ggg gaa ttg cgg gcg gag ctg gag gcg	8352
Gly Ala Asp Ala Pro Gly Ala Gly Glu Leu Arg Ala Glu Leu Glu Ala	
2770 2775 2780	
ctg ggt gct cgg gtg tcg att gtg ccc tgc gac gtg gct gat cgt gac	8400
Leu Gly Ala Arg Val Ser Ile Val Pro Cys Asp Val Ala Asp Arg Asp	
2785 2790 2795 2800	
gca gtg gct gga gtg ttg gca ggg atc ggt ggg gag tgt ccg ctg act	8448
Ala Val Ala Gly Val Leu Ala Gly Ile Gly Gly Glu Cys Pro Leu Thr	
2805 2810 2815	
gcg gtg gta cac gcc gcc ggg gtc ggc gag gcg ggc gac gta gtg gag	8496
Ala Val Val His Ala Ala Gly Val Gly Glu Ala Gly Asp Val Val Glu	
2820 2825 2830	
atg ggt ttg gcg gat ttt gca gcg gtg ttg tcg gcg aag gtg cgt ggt	8544
Met Gly Leu Ala Asp Phe Ala Ala Val Leu Ser Ala Lys Val Arg Gly	
2835 2840 2845	
gcg gcg aat ctg gac gag ttg ctg gcc gac tcg gag ctg gat gcg ttt	8592
Ala Ala Asn Leu Asp Glu Leu Leu Ala Asp Ser Glu Leu Asp Ala Phe	
2850 2855 2860	
gtg atg ttc tcc tcg gtg tcg ggg gtg tgg gga gcc ggc gga cag ggt	8640
Val Met Phe Ser Ser Val Ser Gly Val Trp Gly Ala Gly Gly Gln Gly	
2865 2870 2875 2880	
gcg tat gcg gct gcg aac gcc tac ttg gat gcg ttg gcc gag cag cgt	8688
Ala Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln Arg	
2885 2890 2895	
cgg gcg agg gga ttg gtc ggg acc gcg gtt gcg tgg gga ccg tgg gcc	8736
Arg Ala Arg Gly Leu Val Gly Thr Ala Val Ala Trp Gly Pro Trp Ala	
2900 2905 2910	
ggt gac ggc atg gcc gcc ggc gaa acc ggc gca cag ctg cac cgg atg	8784
Gly Asp Gly Met Ala Ala Gly Glu Thr Gly Ala Gln Leu His Arg Met	
2915 2920 2925	
ggc ctg gcg tcg atg gaa ccg agc gcg gcg ctg ctg gca ctt cag ggt	8832
Gly Leu Ala Ser Met Glu Pro Ser Ala Ala Leu Leu Ala Leu Gln Gly	
2930 2935 2940	

gca ttg gac cgc gat gag acc tcc ctc gtc gtg gcc gat gtc gat tgg 3880  
Ala Leu Asp Arg Asp Glu Thr Ser Leu Val Val Ala Asp Val Asp Trp  
2945 2950 2955 2960

gca cgg ttc gcc cca gcc ttc acc tcg gca cgt cga cgc ccg ctg ctg 8928  
Ala Arg Phe Ala Pro Ala Phe Thr Ser Ala Arg Arg Arg Pro Leu Leu  
2965 2970 2975

gac acc atc gac gag gcc cga gcc gca ttg gaa acc acc ggc gaa caa 8976  
Asp Thr Ile Asp Glu Ala Arg Ala Ala Leu Glu Thr Thr Gly Glu Gln  
2980 2985 2990

gcg ggc aca ggc aaa ccc gtt gag ctg acg caa cgc ctg gcc gga ctg 9024  
Ala Gly Thr Gly Lys Pro Val Glu Leu Thr Gln Arg Leu Ala Gly Leu  
2995 3000 3005

Tcg cgg aag gaa cgc gac gat gcg gta ttg gat ctg gtg cgg gcg gag 9072  
 Ser Arg Lys Glu Arg Asp Asp Ala Val Leu Asp Leu Val Arg Ala Glu  
 3010 3015 3020

acg gcg gct gtg ctg gga cgc gac gat gcc acg gcc ctg gcg cca tcg 9120  
Thr Ala Ala Val Leu Gly Arg Asp Asp Ala Thr Ala Leu Ala Pro Ser  
3025 3030 3035 3040

cgg ccg ttc cag gaa ctc gga ttc gac tcc ttg atg gcg gtg gag ctg 9168  
 Arg Pro Phe Gln Glu Leu Gly Phe Asp Ser Leu Met Ala Val Glu Leu  
 3045 3050 3055

cgc aac cgg ctg aac acc gcc acc ggg atc cag ctg ccc gcc agc acg 9216  
 Arg Asn Arg Leu Asn Thr Ala Thr Gly Ile Gln Leu Pro Ala Ser Thr  
 3060 3065 3070

att ttc gac tac ccc aat gcc gag tcg ctg tcg cgt cac ctc tgc gcc 9264  
Ile Phe Asp Tyr Pro Asn Ala Glu Ser Leu Ser Arg His Leu Cys Ala  
3075 3080 3085

gag ctt ttc cca acg gag act acc gtg gac tcg gcc ctt gcc gag ctc 9312  
Glu Leu Phe Pro Thr Glu Thr Thr Val Asp Ser Ala Leu Ala Glu Leu  
3090 3095 3100

gat cga atc gag cag cag ctc tcg atg ctc acc ggc gaa gcg cgg gca 9360  
Asp Arg Ile Glu Gln Gln Leu Ser Met Leu Thr Gly Glu Ala Arg Ala  
3105 3110 3115 3120

cgg gac cga atc gcg aca cga ctg cga gcc ctc cac gag aag tgg aac 9408  
Arg Asp Arg Ile Ala Thr Arg Leu Arg Ala Leu His Glu Lys Trp Asn  
3125 3130 3135

agc gca gct gaa gta ccg acc gga gcc gat gtc ctg agc acg ctc gat 9456  
 Ser Ala Ala Glu Val Pro Thr Gly Ala Asp Val Leu Ser Thr Leu Asp  
 3140 3145 3150

tcg gcg acg cac gac gag ata ttc gag ttc atc gac aac gag ctc gac 9504  
 Ser Ala Thr His Asp Glu Ile Phe Glu Phe Ile Asp Asn Glu Leu Asp  
 3155 3160 3165

ctg tcc tga 9513  
 Leu Ser  
 3170

<210> 46  
 <211> 3170  
 :212> PRT  
 :213> Saccharopolyspora spinosa

<400> 46  
 Met Ser Asn Glu Glu Lys Leu Arg Glu Tyr Leu Arg Arg Ala Leu Val  
 1 5 10 15

Asp Leu His Gln Ala Arg Glu Arg Leu His Glu Ala Glu Ser Gly Glu  
 20 25 30

Arg Glu Pro Ile Ala Ile Val Ala Met Gly Cys Arg Tyr Pro Gly Gly  
 35 40 45

Val Gln Asp Pro Glu Gly Leu Trp Lys Leu Val Ala Ser Gly Gly Asp  
 50 55 60

Ala Ile Gly Glu Phe Pro Ala Asp Arg Gly Trp His Leu Asp Glu Leu  
 65 70 75 80

Tyr Asp Pro Asp Pro Asp Gln Pro Gly Thr Cys Tyr Thr Arg His Gly  
 85 90 95

Gly Phe Leu His Asp Ala Gly Glu Phe Asp Ala Gly Phe Phe Asp Ile  
 100 105 110

Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
 115 120 125

Glu Ile Ser Trp Glu Thr Val Glu Ser Ala Gly Met Asp Pro Arg Ser  
 130 135 140

Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met Tyr Glu Gly



145		150		155		160
Tyr Asp Thr Gly Ala His Arg Ala Gly Glu Gly Val Glu Gly Tyr Leu						
	165			170		175
Gly Thr Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Ala						
	180			185		190
Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser						
	195			200		205
Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Gln Gly Glu						
	210			215		220
Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro Glu						
	225			230		240
Arg Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg						
	245			250		255
Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly						
	260			265		270
Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly						
	275			280		285
His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly						
	290			295		300
Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Leu Ala Gln Glu Arg Val						
	305			310		315
Ile Gln Gln Val Leu Thr Ser Ala Gly Leu Ser Ala Ser Asp Val Asp						
	325			330		335
Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu						
	340			345		350
Ala Gln Ala Leu Ile Ala Ala Tyr Gly Gln Asp Arg Asp Arg Asp Arg						
	355			360		365
Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala						
	370			375		380
Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala Met Arg His						
	385			390		395
Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro Asn Ser His Val						

405

410

415

Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Glu Asn Ile Arg Trp  
420 425 430

Pro Gly Thr Gly Thr Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser  
435 440 445

Gly Thr Asn Ala His Val Ile Leu Glu His Asp Pro Leu Ala Val Thr  
450 455 460

Glu Asn Glu Glu Ala Ala Gln Ser Pro Ala Pro Gly Ile Val Pro Trp  
465 470 475 480

Ala Leu Ser Gly Arg Ser Ser Thr Ala Leu Arg Ala Gln Ala Glu Arg  
485 490 495

Leu Arg Glu Leu Cys Glu Gln Thr Asp Pro Asp Pro Val Asp Val Gly  
500 505 510

Phe Ser Leu Ala Ala Thr Arg Thr Ala Trp Glu His Arg Ala Val Val  
515 520 525

Leu Gly Arg Asp Ser Ala Thr Leu Arg Ser Gly Leu Gly Val Val Ala  
530 535 540

Ser Gly Glu Pro Ala Val Asp Val Val Glu Gly Ser Val Leu Asp Gly  
545 550 555 560

Glu Val Val Phe Val Phe Pro Gly Gln Gly Trp Gln Trp Ala Gly Met  
565 570 575

Ala Val Asp Leu Leu Asp Ala Ser Pro Thr Phe Ala Arg His Met Asp  
580 585 590

Glu Cys Ala Thr Ala Leu Arg Arg Tyr Val Asp Trp Ser Leu Val Asp  
595 600 605

Val Leu Arg Gly Ala Glu Asn Ser Pro Pro Leu Asp Arg Val Asp Val  
610 615 620

Leu Gln Pro Ala Ser Phe Ala Val Met Val Ser Leu Ala Glu Val Trp  
625 630 635 640

Arg Ser Tyr Gly Val Arg Pro Ala Ala Val Val Gly His Ser Gln Gly  
645 650 655

Glu Ile Ala Ala Ala Cys Ala Ala Gly Val Leu Pro Leu Glu Asp Ala

660	665	670
Ala Arg Leu Val Ala Leu Arg Ser Arg Ala Leu Lys Gly Leu Ser Gly		
675	680	685
Arg Gly Gly Met Ala Ser Leu Ala Cys Pro Ala Asp Glu Val Ala Ala		
690	695	700
Leu Phe Ala Gly Ser Gly Gly Arg Leu Glu Val Ala Ala Ile Asn Gly		
705	710	715 720
Pro Arg Ser Val Val Val Ser Gly Asp Leu Glu Ala Val Asp Glu Leu		
	725	730 735
Leu Ala Glu Cys Ala Glu Lys Asp Met Arg Ala Arg Arg Ile Pro Val		
	740	745 750
Asp Tyr Ala Ser His Ser Ala His Val Glu Val Val Arg Ser Pro Val		
	755	760 765
Leu Ala Ala Ala Ala Gly Val Arg His Arg Asp Gly Gln Val Pro Trp		
	770	775 780
Trp Ser Thr Val Ile Gly Asp Trp Val Asp Pro Ala Arg Leu Asp Gly		
	785	790 795 800
Glu Tyr Trp Tyr Arg Asn Leu Arg Gln Pro Val Arg Phe Glu His Ala		
	805	810 815
Val Gln Gly Leu Val Glu Arg Gly Phe Gly Leu Phe Ile Glu Met Ser		
	820	825 830
Ala His Pro Val Leu Thr Thr Ala Val Glu Glu Thr Gly Ala Glu Ser		
	835	840 845
Glu Thr Ala Val Ala Ala Val Gly Thr Leu Arg Arg Asp Ser Gly Gly		
	850	855 860
Leu Arg Arg Leu Leu His Ser Leu Ala Glu Ala Tyr Val Arg Gly Ala		
	865	870 875 880
Thr Val Asp Trp Ala Val Ala Phe Gly Gly Ala Gly Arg Arg Leu Asp		
	885	890 895
Leu Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Asp Lys Gly		
	900	905 910
Ala Ala Ser Asp Glu Ala Arg Ala Val Ser Asp Pro Ala Ala Gly Trp		

915

920

925

Phe Trp Gln Ala Val Ala Arg Gln Asp Leu Lys Ser Val Ser Asp Ala  
 930 935 940

Leu Asp Leu Asp Ala Asp Ala Pro Leu Ser Ala Thr Leu Pro Ala Leu  
 945 950 955 960

Ser Val Trp His Arg Gln Glu Arg Glu Arg Val Leu Ala Asp Gly Trp  
 965 970 975

Arg Tyr Arg Val Asp Trp Val Arg Val Ala Pro Gln Pro Val Arg Arg  
 980 985 990

Thr Arg Glu Thr Trp Leu Leu Val Val Pro Pro Gly Gly Ile Glu Glu  
 995 1000 1005

Ala Leu Val Glu Arg Leu Thr Asp Ala Leu Asn Thr Arg Gly Ile Ser  
 1010 1015 1020

Thr Leu Arg Leu Asp Val Pro Pro Ala Ala Thr Ser Gly Glu Leu Ala  
 1025 1030 1035 1040

Thr Glu Leu Arg Ala Ala Ala Asp Gly Asp Pro Val Lys Ala Ile Leu  
 1045 1050 1055

Ser Leu Thr Ala Leu Asp Glu Arg Pro His Pro Glu Cys Lys Asp Val  
 1060 1065 1070

Pro Ser Gly Ile Ala Leu Leu Leu Asn Leu Val Lys Ala Leu Gly Glu  
 1075 1080 1085

Ala Asp Leu Arg Ile Pro Leu Trp Thr Ile Thr Arg Gly Ala Val Lys  
 1090 1095 1100

Ala Gly Pro Ala Asp Arg Leu Leu Arg Pro Met Gln Ala Gln Ala Trp  
 1105 1110 1115 1120

Gly Leu Gly Arg Val Ala Ala Leu Glu His Pro Glu Arg Trp Gly Gly  
 1125 1130 1135

Leu Ile Asp Leu Pro Asp Ser Leu Asp Gly Asp Val Leu Thr Arg Leu  
 1140 1145 1150

Gly Glu Ala Leu Thr Asn Gly Leu Ala Glu Asp Gln Leu Ala Ile Arg  
 1155 1160 1165

Gln Ser Gly Val Leu Ala Arg Arg Leu Val Pro Ala Pro Ala Asn Gln

1170	1175	1180
Pro Ala Gly Arg Lys Trp Arg Pro Arg Gly Ser Ala Leu Ile Thr Gly		
185	1190	1195 1200
Gly Leu Gly Ala Val Gly Ala Gln Val Ala Arg Trp Leu Ala Glu Ile		
1205	1210	1215
Gly Ala Glu Arg Ile Val Leu Thr Ser Arg Arg Gly Asn Gln Ala Ala		
1220	1225	1230
Gly Ala Ala Glu Leu Glu Ala Glu Leu Arg Ala Leu Gly Ala Gln Val		
1235	1240	1245
Ser Ile Val Ala Cys Asp Val Thr Asp Arg Ala Glu Met Ser Ala Leu		
1250	1255	1260
Leu Ala Glu Phe Asp Val Thr Ala Val Phe His Ala Ala Gly Val Gly		
265	1270	1275 1280
Arg Leu Leu Pro Leu Ala Glu Thr Asp Gln Asn Gly Leu Ala Glu Ile		
1285	1290	1295
Cys Ala Ala Lys Val Arg Gly Ala Gln Val Leu Asp Glu Leu Cys Asp		
1300	1305	1310
Ser Thr Asp Leu Asp Ala Phe Val Leu Phe Ser Ser Gly Ala Gly Val		
1315	1320	1325
Trp Gly Gly Gly Gly Gln Gly Ala Tyr Gly Ala Ala Asn Ala Phe Leu		
1330	1335	1340
Asp Thr Leu Ala Glu Gln Arg Arg Ala Arg Gly Leu Pro Ala Thr Ser		
1345	1350	1355 1360
Ile Ser Trp Gly Ser Trp Ala Gly Gly Gly Met Ala Asp Gly Ala Ala		
1365	1370	1375
Gly Glu His Leu Arg Arg Arg Gly Ile Arg Pro Met Pro Ala Ala Ser		
1380	1385	1390
Ala Ile Leu Ala Leu Gln Glu Val Leu Asp Gln Asp Glu Thr Cys Val		
1395	1400	1405
Ser Ile Ala Asp Val Asp Trp Asp Arg Phe Val Pro Thr Phe Ala Ala		
1410	1415	1420
Thr Arg Ala Thr Arg Leu Phe Asp Glu Val Pro Ala Ala Arg Lys Ala		

425 1430 1435 1440

Met Pro Ala Asn Gly Pro Ala Glu Pro Gly Gly Ser Pro Phe Ala Arg  
1445 1450 1455

Asn Leu Ala Glu Leu Pro Glu Ala Gln Arg Arg His Glu Leu Val Asp  
1460 1465 1470

Leu Val Cys Ala Gln Val Ala Thr Val Leu Gly His Gly Ser Arg Glu  
1475 1480 1485

Glu Val Gln Pro Glu Arg Ala Phe Arg Ala Leu Gly Phe Asp Ser Leu  
1490 1495 1500

Met Ala Val Asp Leu Arg Asn Arg Leu Thr Thr Ala Thr Gly Leu Arg  
505 1510 1515 1520

Leu Pro Thr Thr Thr Val Phe Asp Tyr Pro Asn Pro Ala Ala Leu Ala  
1525 1530 1535

Ala His Leu Leu Glu Glu Leu Val Gly Asp Val Ala Ser Ala Ala Val  
1540 1545 1550

Thr Ala Ala Ser Ala Pro Ala Ser Asp Glu Pro Ile Ala Ile Val Ala  
1555 1560 1565

Met Ser Cys Arg Phe Pro Gly Gly Ala His Ser Pro Glu Asp Leu Trp  
1570 1575 1580

Arg Leu Val Ala Ala Gly Thr Glu Val Ile Gly Glu Phe Pro Ser Asp  
585 1590 1595 1600

Arg Gly Trp Asp Ala Glu Gly Leu Tyr Asp Pro Asp Ala Ser Arg Pro  
1605 1610 1615

Gly Thr Thr Tyr Ala Arg Met Ala Gly Phe Leu Tyr Asp Ala Gly Glu  
1620 1625 1630

Phe Asp Ala Asp Leu Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met  
1635 1640 1645

Asp Pro Gln Gln Arg Leu Val Leu Glu Ile Ala Trp Glu Ala Leu Glu  
1650 1655 1660

Arg Ala Gly Ile Asp Pro Leu Ser Leu Lys Gly Ser Gly Val Gly Thr  
665 1670 1675 1680

Tyr Ile Gly Ala Gly Ser Arg Gly Tyr Ala Thr Asp Val Arg Gln Phe

1685	1690	1695
Pro Glu Glu Ala Glu Gly Tyr Leu Leu Thr Gly Thr Ser Ala Ser Val		
1700	1705	1710
Leu Ser Gly Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val		
1715	1720	1725
Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala		
1730	1735	1740
Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly		
745	1750	1755 1760
Val Thr Val Met Ser Thr Pro Glu Met Phe Val Glu Phe Ser Arg Gln		
1765	1770	1775
Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Glu Ser Ala		
1780	1785	1790
Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Leu Leu Leu Glu Arg		
1795	1800	1805
Leu Ser Asp Ala His Arg Asn Gly His Arg Val Leu Ala Val Val Arg		
1810	1815	1820
Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Ala Ala Pro		
825	1830	1835 1840
Asn Gly Pro Ser Gln Gln Arg Val Ile Asn Gln Ala Leu Ala Asn Ala		
1845	1850	1855
Ala Leu Ser Ala Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly		
1860	1865	1870
Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr		
1875	1880	1885
Gly Gln Ala Arg Glu Arg Asp Arg Pro Leu Trp Leu Gly Ser Val Lys		
1890	1895	1900
Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile		
905	1910	1915 1920
Lys Met Val Met Ala Met Arg His Gly Gln Leu Pro Ala Ser Leu His		
1925	1930	1935
Ala Asp Glu Pro Thr Ser Glu Val Asp Trp Ser Ser Gly Ala Val Arg		

1940

1945

50

Leu Leu Ala Glu Gln Val Pro Trp Pro Glu Ser Asp Arg Val Arg Arg  
1955 1960 1965

Val Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile  
1970 1975 1980

Leu Glu Gln Ala Thr Asn Ala Pro Asp Ser Thr Ala Glu Thr Asp Lys  
985 1990 1995 2000

Thr Glu Ser Gly Ser Thr Val Asp Ile Pro Val Val Pro Trp Leu Val  
2005 2010 2015

Ser Gly Lys Thr Thr Asp Ser Leu Arg Gly Gln Ala Glu Arg Val Leu  
2020 2025 2030

Ser Gln Val Glu Ser Arg Pro Glu Gln Arg Ser Leu Asp Val Ala Tyr  
2035 2040 2045

Ser Leu Ala Ser Gly Arg Ala Ala Leu Asp Glu Arg Ala Val Val Leu  
2050 2055 2060

Gly Ala Asp Arg Gly Glu Leu Val Ala Gly Leu Ala Ala Leu Ala Ala  
065 2070 2075 2080

Gly Gln Glu Ala Ser Gly Val Ile Ser Gly Thr Arg Ala Ser Ala Arg  
2085 2090 2095

Phe Gly Phe Val Phe Ser Gly Gln Gly Gly Gln Trp Leu Gly Met Gly  
2100 2105 2110

Arg Ala Leu Tyr Ser Lys Phe Pro Val Phe Ala Ala Ala Phe Asp Glu  
2115 2120 2125

Ala Cys Ala Glu Leu Glu Ala His Leu Gly Glu Asp Arg Arg Val Arg  
2130 2135 2140

Asp Val Val Phe Gly Ser Asp Ala Gln Leu Leu Asp Gln Thr Leu Trp  
145 2150 2155 2160

Ala Gln Ser Gly Leu Phe Ala Leu Gln Ala Gly Leu Leu Gly Leu Leu  
2165 2170 2175

Gly Ser Trp Gly Val Arg Pro Asp Val Val Met Gly His Ser Val Gly  
2180 2185 2190

Glu Leu Ala Ala Ala Phe Ala Ala Gly Val Leu Ser Leu Arg Asp Ala



2195	2200	2205
Ala Arg Leu Val Ala Ala Arg Ala Arg Leu Met Gln Ala Leu Pro Ser		
2210	2215	2220
Asp Gly Ala Met Leu Ala Val Ala Ala Gly Glu Asp Leu Val Arg Pro		
2225	2230	2235 2240
Leu Leu Ala Gly Arg Glu Glu Ser Val Ser Val Ala Ala Leu Asn Ala		
2245	2250	2255
Pro Gly Ser Val Val Leu Ser Gly Asp Arg Glu Val Leu Ala Ser Ile		
2260	2265	2270
Val Gly Arg Leu Thr Glu Leu Arg Val Arg Thr Arg Arg Leu Arg Val		
2275	2280	2285
Ser His Ala Phe His Ser His Arg Met Asp Pro Met Leu Gly Glu Phe		
2290	2295	2300
Ala Gln Ile Ala Glu Ser Ala Glu Phe Gly Lys Pro Thr Thr Pro Leu		
305	2310	2315 2320
Val Ser Thr Leu Thr Gly Glu Leu Asp Arg Ala Ala Glu Met Ser Thr		
2325	2330	2335
Pro Gly Tyr Trp Val Arg Gln Ala Arg Glu Pro Val Arg Phe Ala Asp		
2340	2345	2350
Gly Val Gln Ala Leu Ala Ala Gln Gly Ile Gly Thr Val Val Glu Leu		
2355	2360	2365
Gly Pro Asp Gly Thr Leu Ala Ala Leu Val Arg Glu Cys Ala Thr Glu		
2370	2375	2380
Ser Asp Arg Val Gly Arg Ile Ser Ser Ile Pro Leu Met Arg Arg Glu		
385	2390	2395 2400
Arg Asp Glu Thr Arg Ser Val Met Thr Ala Leu Ala His Leu His Thr		
2405	2410	2415
Arg Gly Gly Glu Val Asp Trp Gln Ala Phe Phe Ala Gly Thr Gly Ala		
2420	2425	2430
Arg Gln Leu Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln His Tyr Trp		
2435	2440	2445
Ile Glu Ser Ser Ala Arg Pro Ala Arg Asp Arg Ala Asp Ile Gly Glu		

2450

2455

2460

Val Ala Glu Gln Phe Trp Thr Ala Val Asp Gln Gly Asp Leu Ala Thr  
 465 2470 2475 2480

Leu Val Ala Ala Leu Asp Leu Gly Ala Asp Asp Asp Thr Cys Ala Ser  
 2485 2490 2495

Leu Ser Asp Val Leu Pro Ala Leu Ser Ser Trp Arg Ser Gly Leu Arg  
 2500 2505 2510

Asn Arg Ser Leu Val Asp Ser Cys Arg Tyr Arg Ile Ser Trp His Ser  
 2515 2520 2525

Ser Arg Glu Val Pro Ala Pro Lys Ile Ser Gly Thr Trp Leu Leu Val  
 2530 2535 2540

Val Pro Gly Ala Ala Asp Asp Gly Leu Val Thr Ala Leu Thr Ser Ser  
 545 2550 2555 2560

Leu Val Gly Gly Gly Ala Glu Val Val Arg Ile Gly Leu Ser Glu Glu  
 2565 2570 2575

Asp Pro His Arg Glu Asp Val Ala Gln Arg Leu Ala Asn Ala Leu Thr  
 2580 2585 2590

Asp Ala Gly Gln Leu Gly Gly Val Leu Ser Leu Leu Gly Leu Asp Glu  
 2595 2600 2605

Ser Pro Ala Pro Gly Phe Ser Cys Leu Pro Thr Gly Phe Ala Leu Thr  
 2610 2615 2620

Val Gln Leu Leu Arg Ala Leu Arg Lys Ala Asp Val Glu Ala Pro Phe  
 625 2630 2635 2640

Trp Ala Val Thr Arg Gly Gly Val Ala Leu Glu Asp Val Arg Val Ser  
 2645 2650 2655

Pro Glu Gln Ala Leu Val Trp Gly Leu Leu Arg Val Ala Gly Leu Glu  
 2660 2665 2670

His Pro Glu Phe Trp Gly Gly Leu Ile Asp Leu Pro Ser Asp Trp Asp  
 2675 2680 2685

Asp Arg Leu Gly Ala Arg Leu Ala Gly Val Leu Ala Asp Gly Gly Glu  
 2690 2695 2700

Asp Gln Val Ala Ile Arg Arg Gly Gly Val Phe Val Arg Arg Leu Glu

705

2710

2715

2720

Arg Ala Gly Ala Ser Gly Ala Gly Ser Val Trp Arg Pro Arg Gly Thr  
2725 2730 2735

Val Leu Val Thr Gly Gly Thr Gly Gly Leu Gly Ala His Val Ala Arg  
2740 2745 2750

Trp Leu Ala Gly Ala Gly Ala Glu His Val Val Leu Thr Ser Arg Arg  
2755 2760 2765

Gly Ala Asp Ala Pro Gly Ala Gly Glu Leu Arg Ala Glu Leu Glu Ala  
2770 2775 2780

Leu Gly Ala Arg Val Ser Ile Val Pro Cys Asp Val Ala Asp Arg Asp  
2785 2790 2795 2800

Ala Val Ala Gly Val Leu Ala Gly Ile Gly Gly Glu Cys Pro Leu Thr  
2805 2810 2815

Ala Val Val His Ala Ala Gly Val Gly Glu Ala Gly Asp Val Val Glu  
2820 2825 2830

Met Gly Leu Ala Asp Phe Ala Ala Val Leu Ser Ala Lys Val Arg Gly  
2835 2840 2845

Ala Ala Asn Leu Asp Glu Leu Leu Ala Asp Ser Glu Leu Asp Ala Phe  
2850 2855 2860

Val Met Phe Ser Ser Val Ser Gly Val Trp Gly Ala Gly Gly Gln Gly  
2865 2870 2875 2880

Ala Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Gln Arg  
2885 2890 2895

Arg Ala Arg Gly Leu Val Gly Thr Ala Val Ala Trp Gly Pro Trp Ala  
2900 2905 2910

Gly Asp Gly Met Ala Ala Gly Glu Thr Gly Ala Gln Leu His Arg Met  
2915 2920 2925

Gly Leu Ala Ser Met Glu Pro Ser Ala Ala Leu Leu Ala Leu Gln Gly  
2930 2935 2940

Ala Leu Asp Arg Asp Glu Thr Ser Leu Val Val Ala Asp Val Asp Trp  
2945 2950 2955 2960

Ala Arg Phe Ala Pro Ala Phe Thr Ser Ala Arg Arg Arg Pro Leu Leu

Asp Thr Ile Asp Glu Ala Arg Ala Ala Leu Glu Thr Thr Gly Glu Gln  
 2980 2985 2990

Ala Gly Thr Gly Lys Pro Val Glu Leu Thr Gln Arg Leu Ala Gly Leu  
 2995 3000 3005

Ser Arg Lys Glu Arg Asp Asp Ala Val Leu Asp Leu Val Arg Ala Glu  
 3010 3015 3020

Thr Ala Ala Val Leu Gly Arg Asp Asp Ala Thr Ala Leu Ala Pro Ser  
 025 3030 3035 3040

Arg Pro Phe Gln Glu Leu Gly Phe Asp Ser Leu Met Ala Val Glu Leu  
 3045 3050 3055

Arg Asn Arg Leu Asn Thr Ala Thr Gly Ile Gln Leu Pro Ala Ser Thr  
 3060 3065 3070

Ile Phe Asp Tyr Pro Asn Ala Glu Ser Leu Ser Arg His Leu Cys Ala  
 3075 3080 3085

Glu Leu Phe Pro Thr Glu Thr Thr Val Asp Ser Ala Leu Ala Glu Leu  
 3090 3095 3100

Asp Arg Ile Glu Gln Gln Leu Ser Met Leu Thr Gly Glu Ala Arg Ala  
 105 3110 3115 3120

Arg Asp Arg Ile Ala Thr Arg Leu Arg Ala Leu His Glu Lys Trp Asn  
 3125 3130 3135

Ser Ala Ala Glu Val Pro Thr Gly Ala Asp Val Leu Ser Thr Leu Asp  
 3140 3145 3150

Ser Ala Thr His Asp Glu Ile Phe Glu Phe Ile Asp Asn Glu Leu Asp  
 3155 3160 3165

Leu Ser  
 3170

<210> 47

<211> 14775

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(14772)

<223> ORF21; Polyketidsynthase

<400> 47

atg gcc aat gaa gaa aag ctc ttc ggc tat ctg aag aag gta act gcg 48  
Met Ala Asn Glu Glu Lys Leu Phe Gly Tyr Leu Lys Lys Val Thr Ala  
1 5 10 15

gac ctg cat cag acc cgg cag cgc ctg ctc gcg gcc gag agc cgg agt 96  
Asp Leu His Gln Thr Arg Gln Arg Leu Leu Ala Ala Glu Ser Arg Ser  
20 25 30

cag gag ccg atc gcg atc gtc tcg gcg agc tgc cga ctg ccc ggc ggc 144  
Gln Glu Pro Ile Ala Ile Val Ser Ala Ser Cys Arg Leu Pro Gly Gly  
35 40 45

gtc gac tct ccc gaa gcg ctc tgg caa ctc gtg cgc act ggc acc gac 192  
Val Asp Ser Pro Glu Ala Leu Trp Gln Leu Val Arg Thr Gly Thr Asp  
50 55 60

gcc atc tcg gag ttc ccc gcc gac cgg ggc tgg gat ctc ggc cgg ttg 240  
Ala Ile Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp Leu Gly Arg Leu  
65 70 75 80

tac gat ccc gac ccg aac cac cag gga acg tcg tac acg cgg gcc ggc 288  
Tyr Asp Pro Asp Pro Asn His Gln Gly Thr Ser Tyr Thr Arg Ala Gly  
85 90 95

ggt ttc ctc gca gga gcg ggc gat ttc gac ccc gcc atg ttc ggg att 336  
Gly Phe Leu Ala Gly Ala Gly Asp Phe Asp Pro Ala Met Phe Gly Ile  
100 105 110

tcg ccg cgt gag gcg ttg gcg atg gac ccg cag caa cgg ttg ttg ctg 384  
Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
115 120 125

gag ctg tcc tgg gag gcc ctc gaa cgg gcg ggc ata gac ccg aca tcc 432  
Glu Leu Ser Trp Glu Ala Leu Glu Arg Ala Gly Ile Asp Pro Thr Ser  
130 135 140

ctg cgc ggc agc aag acc ggt gtc ttc ggt ggt gtc acg ccc cag gag 480  
Leu Arg Gly Ser Lys Thr Gly Val Phe Gly Gly Val Thr Pro Gln Glu  
145 150 155 160

tac ggg ccg tcc ttg cag gag atg agc cga aac gct ggg ggt ttt gga 528  
Tyr Gly Pro Ser Leu Gln Glu Met Ser Arg Asn Ala Gly Gly Phe Gly

ctg acc ggg cgg atg gtg agt gtg gcg tcg ggt cgg gtt gcg tat tcg 576  
 Leu Thr Gly Arg Met Val Ser Val Ala Ser Gly Arg Val Ala Tyr Ser  
 180 185 190

ttt ggt ttt gag ggt cct gcg gtg acg gtg gat acg gcg tgt tcg tcg 624  
 Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser  
 195 200 205

tcg ttg gtg gcc ctg cat ttg gcg tgt cag tcg ttg cgt tcc ggc gaa 672  
 Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu  
 210 215 220

tgc gat ctg gcg ctg gcc ggc ggt gtg acg gtg atg gcg aca ccg gcg 720  
 Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Ala  
 225 230 235 240

acg ttc gtg gag ttc tcc cgt cag cgt ggt ttg gct ccg gac ggg cgg 768  
 Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg  
 245 250 255

tgc aag tcg ttc gcg gct gcc gcg gat ggc acc ggg tgg ggt gag ggt 816  
 Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly  
 260 265 270

gcc ggt ctg gtg ttg ctg gag cgg ttg tcg gat gcg cgg cgg aat ggg 864  
 Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly  
 275 280 285

cac gag gtt ctg gcg gtg gtg cgg ggt agc gcg gtg aac cag gac ggc 912  
 His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly  
 290 295 300

gcg tcg aat ggt ttg act gcg ccg aat ggt ccg tcg cag cag cgg gtg 960  
 Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val  
 305 310 315 320

atc acc cag gcg ttg gcg agt gcg ggg ctg tcg gtt tcc gat gtg gat 1008  
 Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp  
 325 330 335

gcg gtc gag gca cat ggg acc ggg acc acg ttg ggt gat ccg atc gag 1056  
 Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu  
 340 345 350

gca cag gcc ctg atc gcc acg tac ggg cag ggc cgg gag aag gat cgg 1104  
 Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly Arg Glu Lys Asp Arg

355	360	365	
ccg ttg tgg ttg ggg tcg gtc aag tcc aac atc ggt cac acg cag gcg	1152		
Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala			
370	375	380	
gcc gct ggc gtt gcc ggc gtc atc aag atg gtc ttg gcg atg cgg cac	1200		
Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Leu Ala Met Arg His			
385	390	395	400
ggg cag ctg ccc gcc acg ttg cat gtg gat gag ccc acg tcg gcg gtg	1248		
Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu Pro Thr Ser Ala Val			
405	410	415	
gac tgg tcg gcg ggt tcg gtc cgg ctt ctc acg gag aac acg ccc tgg	1296		
Asp Trp Ser Ala Gly Ser Val Arg Leu Leu Thr Glu Asn Thr Pro Trp			
420	425	430	
ccg gac agt ggt cgt cct tgc cgg gtg ggg gtg tcg tcg ttc ggg atc	1344		
Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val Ser Ser Phe Gly Ile			
435	440	445	
agc ggc acc aac gca cat gtg att ctc gaa cag tct cca gtc gag cag	1392		
Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ser Pro Val Glu Gln			
450	455	460	
ggc gaa ccg gcc ggg ccg gtc gaa ggc gag cgg gaa ccg gat gta gcc	1440		
Gly Glu Pro Ala Gly Pro Val Glu Gly Glu Arg Glu Pro Asp Val Ala			
465	470	475	480
gtc ccc gtg gtg cct tgg gtg ctg tcg ggt aag aca ccg gag gct gcg	1488		
Val Pro Val Val Pro Trp Val Leu Ser Gly Lys Thr Pro Glu Ala Ala			
485	490	495	
cgg gcg cag gcc gaa cgg gtg cat tcg cat atc gag gac cgg ccg ggg	1536		
Arg Ala Gln Ala Glu Arg Val His Ser His Ile Glu Asp Arg Pro Gly			
500	505	510	
ctg tcg ccg gtg gat gtg gcg tat tcg cta gga atg aca cgc gcg gcg	1584		
Leu Ser Pro Val Asp Val Ala Tyr Ser Leu Gly Met Thr Arg Ala Ala			
515	520	525	
ctg gat gaa cgc gca gtg gtg ttg ggc tcg gac cgt gcc gcg ctc ctg	1632		
Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Ala Ala Leu Leu			
530	535	540	
acc ggg ttg agg gca ttc gcc gac ggc tgc gat gcg ccc gaa gtg gtt	1680		
Thr Gly Leu Arg Ala Phe Ala Asp Gly Cys Asp Ala Pro Glu Val Val			

545	550	555	560	
tcg ggg tct gtg ggg ctt ggt ggc cgc gtc ggg ttc gtg ttc tcg ggt	1728			
Ser Gly Ser Val Gly Leu Gly Gly Arg Val Gly Phe Val Phe Ser Gly				
565	570	575		
cag ggt ggt cag tgg ccg ggg atg ggc cgg ggg ctc tac tcg gtg ttt	1776			
Gln Gly Gly Gln Trp Pro Gly Met Gly Arg Gly Leu Tyr Ser Val Phe				
580	585	590		
ccg gtg ttc gcc gac gcg ttc gac gag gct tgc gcg gag ttg gat gca	1824			
Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala				
595	600	605		
cac ctg ggc cag gaa ctg cgg gtt cgg gat gtg gtg ttc ggt tcg caa	1872			
His Leu Gly Gln Glu Leu Arg Val Arg Asp Val Val Phe Gly Ser Gln				
610	615	620		
gcg tgg ttg ctg gat cgg acg gtg tgg gcg cag tcg ggt ttg ttc gcg	1920			
Ala Trp Leu Leu Asp Arg Thr Val Trp Ala Gln Ser Gly Leu Phe Ala				
625	630	635	640	
ttg cag att ggc ttg ctg cgg ctg ctg ggt tcg tgg ggt gtt cgg ccg	1968			
Leu Gln Ile Gly Leu Leu Arg Leu Leu Gly Ser Trp Gly Val Arg Pro				
645	650	655		
gat gtg gtg ttg ggg cac tcg gtg ggt gag ctg gct gcg gtg cat gcg	2016			
Asp Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Val His Ala				
660	665	670		
gct ggt gtg ttg tcg ttg tcg gag gcc gcg cgg ttg gtg gcg ggt cgc	2064			
Ala Gly Val Leu Ser Leu Ser Glu Ala Ala Arg Leu Val Ala Gly Arg				
675	680	685		
gcc cgg ttg atg cag gcg ttg cct tct ggt ggt gcc atg ctc gcg gtc	2112			
Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala Met Leu Ala Val				
690	695	700		
gct acg ggt gag ttt cag gtc gat cct ctg ctg gat ggg gtg cgg gac	2160			
Ala Thr Gly Glu Phe Gln Val Asp Pro Leu Leu Asp Gly Val Arg Asp				
705	710	715	720	
cgg atc ggt atc gcg gcg gtg aat ggc ccg gaa tcg gtt gtg ctc tct	2208			
Arg Ile Gly Ile Ala Ala Val Asn Gly Pro Glu Ser Val Val Leu Ser				
725	730	735		
ggt gac cgc gag ctg ctc acc gag atc gct gat cgg ttg cac gat cag	2256			
Gly Asp Arg Glu Leu Leu Thr Glu Ile Ala Asp Arg Leu His Asp Gln				



740	745	750	
ggg tgc cgg acc cgg tgg ttg cgg gtg tgc cat gct ttc cat tgc ccc			2304
Gly Cys Arg Thr Arg Trp Leu Arg Val Ser His Ala Phe His Ser Pro			
755	760	765	
cat atg gag ccg atg ctg gag gag ttc gcc cag atc tcc cga ggc cgc			2352
His Met Glu Pro Met Leu Glu Glu Phe Ala Gln Ile Ser Arg Gly Arg			
770	775	780	
gaa tat cac gca ccg gaa ctg ccg atc atc tgc acc ctg atc ggt gag			2400
Glu Tyr His Ala Pro Glu Leu Pro Ile Ile Ser Thr Leu Ile Gly Glu			
785	790	795	800
ctg gac ggt ggt cga gtg atg ggc act ccc gag tac tgg gtg cgt cag			2448
Leu Asp Gly Gly Arg Val Met Gly Thr Pro Glu Tyr Trp Val Arg Gln			
	805	810	815
gtg cgt gag ccc gtc cgt ttc gcc gag ggt gtc cag gcg ctt gtc ggt			2496
Val Arg Glu Pro Val Arg Phe Ala Glu Gly Val Gln Ala Leu Val Gly			
	820	825	830
cag ggt gtc ggc acg att gtc gaa ttg ggt ccg gac ggg gcg ttg tgc			2544
Gln Gly Val Gly Thr Ile Val Glu Leu Gly Pro Asp Gly Ala Leu Ser			
	835	840	845
acg ttg gtc gag gag tgt gtg gcg gaa tcc ggg cgg gtg gcc ggg atc			2592
Thr Leu Val Glu Glu Cys Val Ala Glu Ser Gly Arg Val Ala Gly Ile			
	850	855	860
ccg ctg atg cgc aag gac cgc gac gag gcg cga acc gtg ctg gca gct			2640
Pro Leu Met Arg Lys Asp Arg Asp Glu Ala Arg Thr Val Leu Ala Ala			
	865	870	875
ttg gcg cag atc cac acc cgt ggt ggt gag gtg gac tgg cgg tgc ttt			2688
Leu Ala Gln Ile His Thr Arg Gly Gly Glu Val Asp Trp Arg Ser Phe			
	885	890	895
ttc gcc ggt acc ggg gcg aag caa gtc gac ctg ccc acc tac gcc ttc			2736
Phe Ala Gly Thr Gly Ala Lys Gln Val Asp Leu Pro Thr Tyr Ala Phe			
	900	905	910
cag cgg cag cgg tac tgg ctg gca tcc acc ggg cgt gcg ggt gac gtg			2784
Gln Arg Gln Arg Tyr Trp Leu Ala Ser Thr Gly Arg Ala Gly Asp Val			
	915	920	925
acc gcc gcc gga ttg gcc gag gcg gac cat ccg ctg ctc ggt gcg gtg			2832
Thr Ala Ala Gly Leu Ala Glu Ala Asp His Pro Leu Leu Gly Ala Val			

930	935	940	
gtt gcg ttg gca gac ggc gaa ggt gtg gtg ctg acc ggt cgg ttg aca			2380
Val Ala Leu Ala Asp Gly Glu Gly Val Val Leu Thr Gly Arg Leu Thr			
945	950	955	960
gcg ggt tcg cat ccg tgg ttg tcc gat cac cgg gtg ctg ggc gaa atc			2928
Ala Gly Ser His Pro Trp Leu Ser Asp His Arg Val Leu Gly Glu Ile			
	965	970	975
gtc gtc ccc ggc acc gcg atc gtc gag ctg gtg tgg cac gtc ggc gag			2976
Val Val Pro Gly Thr Ala Ile Val Glu Leu Val Trp His Val Gly Glu			
	980	985	990
cgc ctc ggt tgt ggc cgg gtg gaa gaa ctg gct ttg gaa gcg ccc ctg			3024
Arg Leu Gly Cys Gly Arg Val Glu Glu Leu Ala Leu Glu Ala Pro Leu			
	995	1000	1005
atc ctg ccg gat cat gga gcg gtc cag gtt cag gtg ctg gtg gga ccg			3072
Ile Leu Pro Asp His Gly Ala Val Gln Val Gln Val Leu Val Gly Pro			
	1010	1015	1020
ccc ggg gaa tcc gga gcc cgg tcg gtg gcg ctc tac tcc tgt cct ggc			3120
Pro Gly Glu Ser Gly Ala Arg Ser Val Ala Leu Tyr Ser Cys Pro Gly			
	1025	1030	1035
			1040
gag gcg atc gaa ccc gag tgg aag aag cac gcg acg ggc gtg ctt ctc			3168
Glu Ala Ile Glu Pro Glu Trp Lys Lys His Ala Thr Gly Val Leu Leu			
	1045	1050	1055
cca ccc gtg gcc gcc gag aac cat gag ctg acc gca tgg ccc ccg gag			3216
Pro Pro Val Ala Ala Glu Asn His Glu Leu Thr Ala Trp Pro Pro Glu			
	1060	1065	1070
aat gcg acc gaa atc gat gca gac ggg gtc tac gca ttc ctt gaa ggg			3264
Asn Ala Thr Glu Ile Asp Ala Asp Gly Val Tyr Ala Phe Leu Glu Gly			
	1075	1080	1085
cac ggt ttc gcg tac gga ccg gcc ttt aga tgt ctg cgc ggt gcc tgg			3312
His Gly Phe Ala Tyr Gly Pro Ala Phe Arg Cys Leu Arg Gly Ala Trp			
	1090	1095	1100
cga cga ggc ggg gag gtg ttc gcc gaa gtc gca ttg ccg gat gac atg			3360
Arg Arg Gly Gly Glu Val Phe Ala Glu Val Ala Leu Pro Asp Asp Met			
	1105	1110	1115
			1120
cag gcg ggg gtc gat cga ttc ggc gtc cac ccc gcg ttg ctg gac gcg			3408
Gln Ala Gly Val Asp Arg Phe Gly Val His Pro Ala Leu Leu Asp Ala			

1125	1130	1135	
gtt ctg cat gcc gcc gca gcc gag acg tcg gtg gtc cag agc gaa gcg			3456
Val Leu His Ala Ala Ala Ala Glu Thr Ser Val Val Gln Ser Glu Ala			
1140	1145	1150	
cgg gtg ccg ttc tcg tgg cgt ggg gtg gaa ctt cgc gcc act gaa agc			3504
Arg Val Pro Phe Ser Trp Arg Gly Val Glu Leu Arg Ala Thr Glu Ser			
1155	1160	1165	
gcg gtg gtg cgg gcg cgc ctc tcg ttg act tcg gat gac gaa ctg tcg			3552
Ala Val Val Arg Ala Arg Leu Ser Leu Thr Ser Asp Asp Glu Leu Ser			
1170	1175	1180	
ttg gtc gca gtg gac ccg gct ggc cga ttc gtg gcc acg gtt gat tcg			3600
Leu Val Ala Val Asp Pro Ala Gly Arg Phe Val Ala Thr Val Asp Ser			
1185	1190	1195	1200
ctg gtg acc cga ccg atc tcc cgg cag cag gtg agg tct ggc gcg atc			3648
Leu Val Thr Arg Pro Ile Ser Arg Gln Gln Val Arg Ser Gly Ala Ile			
1205	1210	1215	
ggt gat tgc ctg ttc gag gtg gag tgg cac cgg aag gcg ttg ttg gga			3696
Gly Asp Cys Leu Phe Glu Val Glu Trp His Arg Lys Ala Leu Leu Gly			
1220	1225	1230	
aca acc gcc ggc gac gac ctt gcc atc gtc ggt gac ggt ccc agt tgg			3744
Thr Thr Ala Gly Asp Asp Leu Ala Ile Val Gly Asp Gly Pro Ser Trp			
1235	1240	1245	
ccg gaa tcg gtg cgc gca acc gca cgg ttc gcg acc ctg gat gag ttc			3792
Pro Glu Ser Val Arg Ala Thr Ala Arg Phe Ala Thr Leu Asp Glu Phe			
1250	1255	1260	
cgt gcg gcc gtg gac tcg gac gtt cct gcc ccg ggt tcg gtg ttg gtc			3840
Arg Ala Ala Val Asp Ser Asp Val Pro Ala Pro Gly Ser Val Leu Val			
1265	1270	1275	1280
gca gct atg tcg gcc gaa gag gtc gag ggt gga tcc ctg ccg tcg cgc			3888
Ala Ala Met Ser Ala Glu Glu Val Glu Gly Gly Ser Leu Pro Ser Arg			
1285	1290	1295	
gcc caa gag tcg acc tcc gat ctg ctg gct ctc gtg cag tcg tgg ctt			3936
Ala Gln Glu Ser Thr Ser Asp Leu Leu Ala Leu Val Gln Ser Trp Leu			
1300	1305	1310	
gcg gac gag ccg ttc gcc gaa tcc cag ctc gtg gtc gtc acg cgt gca			3984
Ala Asp Glu Arg Phe Ala Glu Ser Gln Leu Val Val Val Thr Arg Ala			

1315	1320	1325	
gcg gtg tcg gcc gac tcg gat tcg gac gtc gcg gac ctg gtg ggt gcg	4032		
Ala Val Ser Ala Asp Ser Asp Ser Asp Val Ala Asp Leu Val Gly Ala			
1330	1335	1340	
tcg tcg tgg ggg ttg ttg agt tca gcc cag tcg gag aac ccg ggt cgc	4080		
Ser Ser Trp Gly Leu Leu Ser Ser Ala Gln Ser Glu Asn Pro Gly Arg			
1345	1350	1355	1360
ttc gtg ctg gtg gac gtg gac ggc aca cct gag tcg tgg cag gcg ttg	4128		
Phe Val Leu Val Asp Val Asp Gly Thr Pro Glu Ser Trp Gln Ala Leu			
1365	1370	1375	
ccg gcc gcc gtg cga gca gga gaa ccg cag ctg gca ctt ccg cgc ggc	4176		
Pro Ala Ala Val Arg Ala Gly Glu Pro Gln Leu Ala Leu Arg Arg Gly			
1380	1385	1390	
gtg gcg ctg gtg cct ccg ttg gcg cga ctc acg gtg cgc gag gag ggc	4224		
Val Ala Leu Val Pro Arg Leu Ala Arg Leu Thr Val Arg Glu Glu Gly			
1395	1400	1405	
tcc tcc ccg caa ctc gac acg gac ggg acc gtc ctc atc acg ggt ggc	4272		
Ser Ser Pro Gln Leu Asp Thr Asp Gly Thr Val Leu Ile Thr Gly Gly			
1410	1415	1420	
acc ggt gcg ttg ggg gga gtg gtt gcc cgt cac ctg gtg gag gag cac	4320		
Thr Gly Ala Leu Gly Gly Val Val Ala Arg His Leu Val Glu Glu His			
1425	1430	1435	1440
ggg att ccg cgt ttg gtg ttg gca ggc ccg cgt ggc tgg aat gcg cct	4368		
Gly Ile Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp Asn Ala Pro			
1445	1450	1455	
gga gtc cac gag ttg gtg gat gag ctg gcg cgc gcg ggc gcc gtg gtt	4416		
Gly Val His Glu Leu Val Asp Glu Leu Ala Arg Ala Gly Ala Val Val			
1460	1465	1470	
gag gtg gtg gct tgc gat gtg gct gac cgc acc gat ctg gag cac gtg	4464		
Glu Val Val Ala Cys Asp Val Ala Asp Arg Thr Asp Leu Glu His Val			
1475	1480	1485	
ctg gcc gcc att ccg gtc gac tgg ccg ctg ccg ggg atc gtg cat acc	4512		
Leu Ala Ala Ile Pro Val Asp Trp Pro Leu Arg Gly Ile Val His Thr			
1490	1495	1500	
gct ggg gtg ctg gcc gac gga gtg atc ggg tcc ttg tcg gcg gcg gat	4560		
Ala Gly Val Leu Ala Asp Gly Val Ile Gly Ser Leu Ser Ala Ala Asp			

1505	1510	1515	1520	
gtg ggc acg gtg ttt gcc ccg aag gtg acg ggg gca tgg cat ctg cac				4608
Val Gly Thr Val Phe Ala Pro Lys Val Thr Gly Ala Trp His Leu His				
1525	1530	1535		
gag ttg acc cgc gat ctg gat ctg tcg ttc ttc gtt ctt ttc tct tcc				4656
Glu Leu Thr Arg Asp Leu Asp Leu Ser Phe Phe Val Leu Phe Ser Ser				
1540	1545	1550		
ttc tcc ggg att gcg ggt gcc gca ggg cag gcc aac tac gcg gcg gcg				4704
Phe Ser Gly Ile Ala Gly Ala Ala Gly Gln Ala Asn Tyr Ala Ala Ala				
1555	1560	1565		
aac acg ttc ctg gat gca ttg gcg cgt tat cgc cgg gcg cgt ggg ctg				4752
Asn Thr Phe Leu Asp Ala Leu Ala Arg Tyr Arg Arg Ala Arg Gly Leu				
1570	1575	1580		
cct ggg ttg tcg ttg gcg tgg gga ctg tgg gcg caa ccc agc ggt atg				4800
Pro Gly Leu Ser Leu Ala Trp Gly Leu Trp Ala Gln Pro Ser Gly Met				
1585	1590	1595	1600	
acg agt ggc ttg gac gcg gcg tcg gtg gag cgg ttg gcg cgg acg ggc				4848
Thr Ser Gly Leu Asp Ala Ala Ser Val Glu Arg Leu Ala Arg Thr Gly				
1605	1610	1615		
atc gca gaa ctt tcc acg gag gat gga ctc cgc ctg ttc gat gcc gcg				4896
Ile Ala Glu Leu Ser Thr Glu Asp Gly Leu Arg Leu Phe Asp Ala Ala				
1620	1625	1630		
ttc gcg aag gac cgg gct tgc gtc gtt gcc gct cga ttg gac agg gcg				4944
Phe Ala Lys Asp Arg Ala Cys Val Val Ala Ala Arg Leu Asp Arg Ala				
1635	1640	1645		
ctg ctg gtc ggg aac gga cga tcg cac gcg att ccg gcg ctg ttg agc				4992
Leu Leu Val Gly Asn Gly Arg Ser His Ala Ile Pro Ala Leu Leu Ser				
1650	1655	1660		
gcg ttg gtt cct gtt cgc ggc ggt gtg gcg agg aaa aca gcc aat tct				5040
Ala Leu Val Pro Val Arg Gly Gly Val Ala Arg Lys Thr Ala Asn Ser				
1665	1670	1675	1680	
cag gcc gcg gat gag gac gca ctg ttg ggt ttg gtg cgg gag cac gtt				5088
Gln Ala Ala Asp Glu Asp Ala Leu Leu Gly Leu Val Arg Glu His Val				
1685	1690	1695		
tcg gcc gtg ctg ggt tat tcg ggt gcg gtc gag gtt ggg ggc gac cgt				5136
Ser Ala Val Leu Gly Tyr Ser Gly Ala Val Glu Val Gly Gly Asp Arg				

1700

1705

1710

gct ttc cgt gat ctg ggt ttt gat tcg ttg tct ggc gtg gag ttg cgg 5184  
 Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg  
 1715 1720 1725

aac cgc ctt gcc ggg gtg ctg ggg gtg cgg ttg ccg gcg act gcg gtg 5232  
 Asn Arg Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val  
 1730 1735 1740

ttc gac tat ccg acg ccg cgg gcg ctg gcg cgt ttc ctg cat cag gaa 5280  
 Phe Asp Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu  
 1745 1750 1755 1760

ctg gca ggc gag gtc gcg tcc acg tcg acg ccg gtg acc agg gca gcg 5328  
 Leu Ala Gly Glu Val Ala Ser Thr Ser Thr Pro Val Thr Arg Ala Ala  
 1765 1770 1775

agt gcc gaa gag gat ctt gtt gcg att gtc ggg atg gga tgt cgt ttt 5376  
 Ser Ala Glu Glu Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe  
 1780 1785 1790

ccg ggt ggg gtg tcg tcg ccg gag gag ctt tgg cgg ctg gtg gcc ggc 5424  
 Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly  
 1795 1800 1805

ggc gtg gat gcg gtg gct ggg ttc cca gac gat cgc gcc tgg gat ctc 5472  
 Gly Val Asp Ala Val Ala Gly Phe Pro Asp Asp Arg Gly Trp Asp Leu  
 1810 1815 1820

gcg gcg ttg tac gat cct gat ccc gat cgt ctc ggg acc tcg tat gtg 5520  
 Ala Ala Leu Tyr Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr Val  
 1825 1830 1835 1840

tgt gag ggc ggg ttt ctg cgg gac gcg gcg gag ttc gat gct gac atg 5568  
 Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Asp Met  
 1845 1850 1855

ttc gcc atc agc ccg cgt gag gcg ttg gcg atg gat ccg cag cag cgg 5616  
 Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg  
 1860 1865 1870

ttg ctg ctg gag gtc gcc tgg gaa acc ttg gag cgg gct ggg atc gat 5664  
 Leu Leu Leu Glu Val Ala Trp Glu Thr Leu Glu Arg Ala Gly Ile Asp  
 1875 1880 1885

ccg ttc tcg ttg cac gcc agc cgg acc ggt gtg ttc gcg gcc ttg atg 5712  
 Pro Phe Ser Leu His Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met

1890	1895	1900	
tac cac gac tat ggg gcc cga ttc att acc aga gca ccg gag ggc ttc			5760
Tyr His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly Phe			
1905	1910	1915	1920
gaa ggg cac ctc ggg acg ggc aat gcg ggg agc gtg ctg tcg ggt cgg			5808
Glu Gly His Leu Gly Thr Gly Asn Ala Gly Ser Val Leu Ser Gly Arg			
1925	1930	1935	
gtt gcg tat tcg ttt ggt ttc gag ggt cct gcg gtg acg gtg gat acg			5856
Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr			
1940	1945	1950	
gcg tgt tcg tcg tcg ttg gtg gcg tta cac ctg gcg ggt caa gca ctg			5904
Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu			
1955	1960	1965	
cgg gcc ggt gag tgc gaa ttc gcc ctt gcc ggt ggc gtc acg gtg atg			5952
Arg Ala Gly Glu Cys Glu Phe Ala Leu Ala Gly Gly Val Thr Val Met			
1970	1975	1980	
tcg acg ccg acg acg ttc gtg gag ttc tcc cgt caa cgg ggt ctg gct			6000
Ser Thr Pro Thr Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala			
1985	1990	1995	2000
ccg gat ggg cgg tgc aag tcg ttc gcg gcg gcc gcg gat ggc acc ggg			6048
Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly			
2005	2010	2015	
tgg ggc gag ggt gcc ggt ctg gtg ttg ctg gag cgg ttg tcg gat gcc			6096
Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala			
2020	2025	2030	
cgg cgc aat ggg cac gag gtt ctg gcg gtg gtg cgg ggt agc gcg gtg			6144
Arg Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val			
2035	2040	2045	
aac cag gac ggc gcg tcg aat ggc ttg act gcg cca aat ggt ccg tca			6192
Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser			
2050	2055	2060	
cag caa agg gtg atc acc cag gca ctc acg agt gcc ggg ctg tcc gtg			6240
Gln Gln Arg Val Ile Thr Gln Ala Leu Thr Ser Ala Gly Leu Ser Val			
2065	2070	2075	2080
tcc gac gtg gat gct gtg gag gcg cat ggg acg ggc acg cgg ctt ggt			6288
Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly			

2088

2090

2095

gat ccg atc gag gcg cag gcg ttg atc gct acg tac ggc cgg gat cgt 6336  
 Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg  
 2100 2105 2110

gat ccc ggt cgg ccg ttg tgg ctg ggg tcg gtg aag tcg aat att ggt 6384  
 Asp Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly  
 2115 2120 2125

cac acc cag gcg gcg gcg ggt gtc gct ggt gtg atc aag atg gtg atg 6432  
 His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met  
 2130 2135 2140

gcg atg cgg cag ggg gag ctg ccg cgc acg ttg cac gtg gac gag ccc 6480  
 Ala Met Arg Gln Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro  
 2145 2150 2155 2160

tcc gcg cag gtg gac tgg tct gcg ggc acg gtc caa ctc ctc acg gag 6528  
 Ser Ala Gln Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu  
 2165 2170 2175

aac acg ccc tgg ccc gac agc ggt cgt ctt cgc cgg gcg ggc gtg tca 6576  
 Asn Thr Pro Trp Pro Asp Ser Gly Arg Leu Arg Arg Ala Gly Val Ser  
 2180 2185 2190

tcg ttc ggg atc agt ggc acc aac gcg cac ctg atc ctt gaa caa cct 6624  
 Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu Ile Leu Glu Gln Pro  
 2195 2200 2205

ccg cga gag tcg cag cgc tca aca gag ccg gat tcg ggt tct gtc cgc 6672  
 Pro Arg Glu Ser Gln Arg Ser Thr Glu Pro Asp Ser Gly Ser Val Arg  
 2210 2215 2220

gat ttt ccg gtg gtg ccg tgg atg gtg tcg ggc aaa aca ccc gaa gcg 6720  
 Asp Phe Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala  
 2225 2230 2235 2240

cta tcc gcc cag gca gat gca ttg atg tcc tac ttg agc aat cgc gtt 6768  
 Leu Ser Ala Gln Ala Asp Ala Leu Met Ser Tyr Leu Ser Asn Arg Val  
 2245 2250 2255

gat gct tcc ccg cga gat atc ggt tat tcg ctt gcg gtg acc cgt ccg 6816  
 Asp Ala Ser Pro Arg Asp Ile Gly Tyr Ser Leu Ala Val Thr Arg Pro  
 2260 2265 2270

gcg ttg gac cac cgc gct gtc gtg ctg ggt gcg gat cgt gcc gcg ttg 6864  
 Ala Leu Asp His Arg Ala Val Val Leu Gly Ala Asp Arg Ala Ala Leu



2275	2280	2285	
ctg ccg ggc ttg aaa gcg ctg gcc gtt agt aat gac gct gcc gag gtg			6912
Leu Pro Gly Leu Lys Ala Leu Ala Val Ser Asn Asp Ala Ala Glu Val			
2290	2295	2300	
atc acc ggc act cgt gcc gct ggg ccg gtc gga ttc gtg ttc tcc ggt			6960
Ile Thr Gly Thr Arg Ala Ala Gly Pro Val Gly Phe Val Phe Ser Gly			
2305	2310	2315	2320
caa ggt ggt cag tgg ccc ggg atg gga agc ggg ctc cac tcg gcg ttt			7008
Gln Gly Gly Gln Trp Pro Gly Met Gly Ser Gly Leu His Ser Ala Phe			
2325	2330	2335	
ccg gtg ttc gcc gac gcg ttt gac gaa gcc tgc tgc gag ctg gat gcg			7056
Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Cys Glu Leu Asp Ala			
2340	2345	2350	
cat ctc ggg cag atg gcc cgg cta cga gat gtg ttg tcc ggt tcg gat			7104
His Leu Gly Gln Met Ala Arg Leu Arg Asp Val Leu Ser Gly Ser Asp			
2355	2360	2365	
acg caa ctt ctg gac cag acc ttg tgg gcg cag ccg ggc ctg ttc gcg			7152
Thr Gln Leu Leu Asp Gln Thr Leu Trp Ala Gln Pro Gly Leu Phe Ala			
2370	2375	2380	
ttg caa gtc gga ctc tgg gag ttg ttg ggt tcg tgg ggt gtc cgg ccc			7200
Leu Gln Val Gly Leu Trp Glu Leu Leu Gly Ser Trp Gly Val Arg Pro			
2385	2390	2395	2400
gct gtg gtg ctg ggc cac tcg gtc ggt gag ctg gcg gcg gcg ttc gcg			7248
Ala Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala			
2405	2410	2415	
gct gga gtg ttg tcg ttg cgg gat gcg gct cgg ctg gtg gcg ggc cgt			7296
Ala Gly Val Leu Ser Leu Arg Asp Ala Ala Arg Leu Val Ala Gly Arg			
2420	2425	2430	
gcc cgg ttg atg caa gcc ctg cca act ggc ggt gcc atg ctc gct gcg			7344
Ala Arg Leu Met Gln Ala Leu Pro Thr Gly Gly Ala Met Leu Ala Ala			
2435	2440	2445	
gct gct gga gag gag cag ctg cgc ccg ttg ctg gcc gac tgc ggt gat			7392
Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala Asp Cys Gly Asp			
2450	2455	2460	
cgt gtg ggg atc gcc gcg gtc aac gct ccc ggg tcg gtg gtg ctc tcc			7440
Arg Val Gly Ile Ala Ala Val Asn Ala Pro Gly Ser Val Val Leu Ser			

2465	170	2475	2480	
ggt gat cgg gat gtg ctc gat gac att gcc ggt cgg ctg gac ggg caa				7488
Gly Asp Arg Asp Val Leu Asp Asp Ile Ala Gly Arg Leu Asp Gly Gln				
2485		2490	2495	
ggg atc cgg tcc agg tgg ttg cgg gtt tcg cat gcg ttt cat tcg cat				7536
Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser His				
2500		2505	2510	
cgg atg gat ccg atg ctg gcg gag ttc acc gaa atc gcc cgg agc gtg				7584
Arg Met Asp Pro Met Leu Ala Glu Phe Thr Glu Ile Ala Arg Ser Val				
2515		2520	2525	
gac tac cgg tcg tca ggg ctg ccg atc gtg tcg acg ttg acg ggt gag				7632
Asp Tyr Arg Ser Ser Gly Leu Pro Ile Val Ser Thr Leu Thr Gly Glu				
2530		2535	2540	
ctc gat gag gtc ggc atg ccg gct acg ccg gag tat tgg gtg cgc cag				7680
Leu Asp Glu Val Gly Met Pro Ala Thr Pro Glu Tyr Trp Val Arg Gln				
2545		2550	2555	2560
gtg cga gaa ccc gtc cgc ttc gcc gac ggt gtt gct gcg ctc gcg gct				7728
Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Ala Ala Leu Ala Ala				
2565		2570	2575	
cac ggt gtg agc acc gtc gtc gag gtc ggt ccg gat ggg gtg ttg tcg				7776
His Gly Val Ser Thr Val Val Glu Val Gly Pro Asp Gly Val Leu Ser				
2580		2585	2590	
gcg ctg gtg cag gag tgc gcg gcc gga tcc gat cag ggc gga cgg gtg				7824
Ala Leu Val Gln Glu Cys Ala Ala Gly Ser Asp Gln Gly Gly Arg Val				
2595		2600	2605	
gcc gcg gtt ccg ctc atg cgc agc aat cgc gac gag gcg cac acg gtg				7872
Ala Ala Val Pro Leu Met Arg Ser Asn Arg Asp Glu Ala His Thr Val				
2610		2615	2620	
aca acg gca ttg gcg cag atc cat gtg cgt ggt gct gag gtg gac tgg				7920
Thr Thr Ala Leu Ala Gln Ile His Val Arg Gly Ala Glu Val Asp Trp				
2625		2630	2635	2640
cgg tcg ttt ttc gcc ggt acc ggg gca aag cag gtc gag ctg ccc acg				7968
Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Val Glu Leu Pro Thr				
2645		2650	2655	
tat gcc ttc caa cga cag cgg tac tgg ctt gac tca cca tcc gaa ccg				8016
Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Asp Ser Pro Ser Glu Pro				

2660	2665	2670	
gtc ggg caa tcc gcc gat ccc gcg cgc cag tcg ggc ttc tgg gaa ctc			8064
Val Gly Gln Ser Ala Asp Pro Ala Arg Gln Ser Gly Phe Trp Glu Leu			
2675	2680	2685	
gtc gag cag gaa gat gtc agc gcg ctc agc gcc gct ctg cac att acc			8112
Val Glu Gln Glu Asp Val Ser Ala Leu Ser Ala Ala Leu His Ile Thr			
2690	2695	2700	
ggc gat cac gac gtg cag gcg tcc ctg gaa tcg gtg gtt ccg gtc ctc			8160
Gly Asp His Asp Val Gln Ala Ser Leu Glu Ser Val Val Pro Val Leu			
2705	2710	2715	2720
tcc tcc tgg cat cgc cgg atc cgc aac gaa tcc ctg gtg cac cag tgg			8208
Ser Ser Trp His Arg Arg Ile Arg Asn Glu Ser Leu Val His Gln Trp			
2725	2730	2735	
cgg tac cgg att tcc tgg cat gag cgg gca gat ttg cca gac ccc tcg			8256
Arg Tyr Arg Ile Ser Trp His Glu Arg Ala Asp Leu Pro Asp Pro Ser			
2740	2745	2750	
ttg tcg ggg aca tgg ctc gtc gtc gtg ccg gag ggg tgg tcg gcg agt			8304
Leu Ser Gly Thr Trp Leu Val Val Val Pro Glu Gly Trp Ser Ala Ser			
2755	2760	2765	
cgg caa gtt ctg cgt ttc aac gag atg ttc gag gaa cgg ggt tgc ccg			8352
Arg Gln Val Leu Arg Phe Asn Glu Met Phe Glu Glu Arg Gly Cys Pro			
2770	2775	2780	
gca gtt ctg ttc gag ctc gcc ggg cac gac gag gaa gcc ctg gcg caa			8400
Ala Val Leu Phe Glu Leu Ala Gly His Asp Glu Glu Ala Leu Ala Gln			
2785	2790	2795	2800
cga ttc cgc tcg ttg cct gtt gcg tca ggg gga ata agc ggc gtg ttg			8448
Arg Phe Arg Ser Leu Pro Val Ala Ser Gly Gly Ile Ser Gly Val Leu			
2805	2810	2815	
tcc ttg ctg gcg ctg gat gaa tcg ccg tcc tcg ccg aac gct gct ttg			8496
Ser Leu Leu Ala Leu Asp Glu Ser Pro Ser Ser Pro Asn Ala Ala Leu			
2820	2825	2830	
ccg aat ggc gcg ctg aac tcg ttg gta ctg ctg cga gct ctg ccg gcc			8544
Pro Asn Gly Ala Leu Asn Ser Leu Val Leu Leu Arg Ala Leu Arg Ala			
2835	2840	2845	
gcg gat gtg tcg gcg cca ttg tgg ttg gcg acg tgt ggt ggt gtc gcg			8592
Ala Asp Val Ser Ala Pro Leu Trp Leu Ala Thr Cys Gly Gly Val Ala			

2850	2855	2860	
gtc ggg gat gtg ccg gtg aac ccg ggg cag gcg ctg gtg tgg gga ctg			8640
Val Gly Asp Val Pro Val Asn Pro Gly Gln Ala Leu Val Trp Gly Leu			
2865	2870	2875	2880
ggt cgc gtc gtc ggt ctg gag cat ccg gcc tgg tgg ggt ggc ctg gtc			8688
Gly Arg Val Val Gly Leu Glu His Pro Ala Trp Trp Gly Gly Leu Val			
2885	2890	2895	
gac gtg ccg tgc ttg ctc gat gag gac gct cga gaa cgc ttg tcg gtc			8736
Asp Val Pro Cys Leu Leu Asp Glu Asp Ala Arg Glu Arg Leu Ser Val			
2900	2905	2910	
gtg ttg gca ggt ctt ggc gag gac gag atc gcg gta cgt ccc ggt ggt			8784
Val Leu Ala Gly Leu Gly Glu Asp Glu Ile Ala Val Arg Pro Gly Gly			
2915	2920	2925	
gtg ttc gtg cgg cgg ttg gaa cgc gct ggt gcg gcg tcg ggt gcc ggg			8832
Val Phe Val Arg Arg Leu Glu Arg Ala Gly Ala Ala Ser Gly Ala Gly			
2930	2935	2940	
tcg gtg tgg cgt cct cgg ggg acg gtg ttg gtg acg ggt ggt acg ggc			8880
Ser Val Trp Arg Pro Arg Gly Thr Val Leu Val Thr Gly Gly Thr Gly			
2945	2950	2955	2960
ggt ttg ggg gcg cat gtt gcc cgg tgg ttg gcg ggt gcc ggg gct gag			8928
Gly Leu Gly Ala His Val Ala Arg Trp Leu Ala Gly Ala Gly Ala Glu			
2965	2970	2975	
cat gtg gtg ttg acc agc cgt cga ggc gcg gcg gct ccg ggc gct gga			8976
His Val Val Leu Thr Ser Arg Arg Gly Ala Ala Ala Pro Gly Ala Gly			
2980	2985	2990	
gat ttg cgg gcg gag ctg gag gcg ctg ggc gct cgg gtt tcg atc acg			9024
Asp Leu Arg Ala Glu Leu Glu Ala Leu Gly Ala Arg Val Ser Ile Thr			
2995	3000	3005	
gcc tgc gac gtg gcc gat cgt gac gct ttg gcc gaa gtg ttg gcg acc			9072
Ala Cys Asp Val Ala Asp Arg Asp Ala Leu Ala Glu Val Leu Ala Thr			
3010	3015	3020	
att ccg gat gat tgc ccg ctg acc gcg gtg atg cat gcg gcg ggg gtc			9120
Ile Pro Asp Asp Cys Pro Leu Thr Ala Val Met His Ala Ala Gly Val			
3025	3030	3035	3040
ggt gaa gtc ggc gac gtg gcg tcg atg tgt ttg acc gac ttc gtt ggg			9168
Val Glu Val Gly Asp Val Ala Ser Met Cys Leu Thr Asp Phe Val Gly			

3045

3050

3055

gtg ctg tcg gcg aag gca ggt ggt gcg gcg aat ctc gat gag ttg ctc 9216  
 Val Leu Ser Ala Lys Ala Gly Gly Ala Ala Asn Leu Asp Glu Leu Leu  
 3060 3065 3070

gcc gat gtc gag ctg gat gcc ttc gtg ctg ttc tca tcc gtc tcg ggt 9264  
 Ala Asp Val Glu Leu Asp Ala Phe Val Leu Phe Ser Ser Val Ser Gly  
 3075 3080 3085

gtg tgg ggt gct ggc ggg cag ggc gct tat gcg gcg gcg aat gcc tac 9312  
 Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ala Ala Asn Ala Tyr  
 3090 3095 3100

ttg gat gcg ttg gcg cag cag cgt cgg gca agg ggg ttg gtg ggg act 9360  
 Leu Asp Ala Leu Ala Gln Gln Arg Arg Ala Arg Gly Leu Val Gly Thr  
 3105 3110 3115 3120

gcg gtt gcg tgg ggc ccg tgg gcc ggt gac gga atg gcc gca ggt gaa 9408  
 Ala Val Ala Trp Gly Pro Trp Ala Gly Asp Gly Met Ala Ala Gly Glu  
 3125 3130 3135

ggc ggt gca cag ctg cgc cgg gcc ggc ctg gtg cca atg gct gcg gat 9456  
 Gly Gly Ala Gln Leu Arg Arg Ala Gly Leu Val Pro Met Ala Ala Asp  
 3140 3145 3150

cgg gcg ttg ctg gca ctt cag ggc gca ttg gat cgt gac gag aca tcc 9504  
 Arg Ala Leu Leu Ala Leu Gln Gly Ala Leu Asp Arg Asp Glu Thr Ser  
 3155 3160 3165

ctg gtc gtg gcc gat atg gcg tgg gag agg ttc gcc ccg gtg ttc gcc 9552  
 Leu Val Val Ala Asp Met Ala Trp Glu Arg Phe Ala Pro Val Phe Ala  
 3170 3175 3180

atg tcc cgt cgg cgt ccg ctg ctc gac gag ctg ccc gaa gca cag cag 9600  
 Met Ser Arg Arg Arg Pro Leu Leu Asp Glu Leu Pro Glu Ala Gln Gln  
 3185 3190 3195 3200

gcg ttg gcg gat gcg gag aac acc act gat gct gcg gac tcg gcc gtc 9648  
 Ala Leu Ala Asp Ala Glu Asn Thr Thr Asp Ala Ala Asp Ser Ala Val  
 3205 3210 3215

ccg cta ccg cgg ctc gcg ggc atg gca gcc gcc gaa cgc cgc cgc gcg 9696  
 Pro Leu Pro Arg Leu Ala Gly Met Ala Ala Ala Glu Arg Arg Arg Ala  
 3220 3225 3230

atg ctg gac ctg gtg ctg gcg gag gcc tcg att gtg ttg gga cac aac 9744  
 Met Leu Asp Leu Val Leu Ala Glu Ala Ser Ile Val Leu Gly His Asn

3235

3240

32

ggg tct gac cca gtt ggt ccc gac cgg gcg ttc cag gag ctc gga ttt 9792  
 Gly Ser Asp Pro Val Gly Pro Asp Arg Ala Phe Gln Glu Leu Gly Phe  
 3250 3255 3260

gat tcg ctg atg gcc gtc gaa ctg cgc aac agg ttg ggc gag gca aca 9840  
 Asp Ser Leu Met Ala Val Glu Leu Arg Asn Arg Leu Gly Glu Ala Thr  
 3265 3270 3275 3280

gga ttg agt ctg ccg gcc acg ttg atc ttc gat tat ccg agc cca tcc 9888  
 Gly Leu Ser Leu Pro Ala Thr Leu Ile Phe Asp Tyr Pro Ser Pro Ser  
 3285 3290 3295

gcg ctg gct gag cag ctg gtc ggc gag ctg gtg gga gcg cag ccc gcg 9936  
 Ala Leu Ala Glu Gln Leu Val Gly Glu Leu Val Gly Ala Gln Pro Ala  
 3300 3305 3310

acc acc gtc gtg gcc ggg gcc gat cca gtg gat gat ccg gtt gtc gtg 9984  
 Thr Thr Val Val Ala Gly Ala Asp Pro Val Asp Asp Pro Val Val Val  
 3315 3320 3325

gtc gcg atg gga tgc cgg tat ccg gcc gac gtc tgc tcg ccc gag gag 10032  
 Val Ala Met Gly Cys Arg Tyr Pro Gly Asp Val Cys Ser Pro Glu Glu  
 3330 3335 3340

ctg tgg cag ctg gtt tct gcg gga cgt gat gcg gta tcg acg ttc ccc 10080  
 Leu Trp Gln Leu Val Ser Ala Gly Arg Asp Ala Val Ser Thr Phe Pro  
 3345 3350 3355 3360

gtc gat cgg ggt tgg gac tgc aac acg ttg ttc gac ccg gat ccg gat 10128  
 Val Asp Arg Gly Trp Asp Cys Asn Thr Leu Phe Asp Pro Asp Pro Asp  
 3365 3370 3375

cgg gca ggc agt acc tat gtg cga gaa ggt gcc ttc ctg acc ggt gct 10176  
 Arg Ala Gly Ser Thr Tyr Val Arg Glu Gly Ala Phe Leu Thr Gly Ala  
 3380 3385 3390

gat cgg ttc gac gcc ggg ttc ttc gcc atc agc cct cgc gag gcg cgc 10224  
 Asp Arg Phe Asp Ala Gly Phe Phe Gly Ile Ser Pro Arg Glu Ala Arg  
 3395 3400 3405

gca atg gat ccg cag cag agg ttg ttg ctc gaa gtg gcg tgg gag gtt 10272  
 Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Val Ala Trp Glu Val  
 3410 3415 3420

ttc gaa cga gca gga atc gct ccg ctg tcg ttg cgg ggt agc agg acc 10320  
 Phe Glu Arg Ala Gly Ile Ala Pro Leu Ser Leu Arg Gly Ser Arg Thr

3425	3430	3435	3440	
ggt gtg ttc gcg ggg acc aat ggg cag gac cac ggt gcg aaa gtg gct				10368
Gly Val Phe Ala Gly Thr Asn Gly Gln Asp His Gly Ala Lys Val Ala				
3445	3450	3455		
gcc gcg ccg gag gcg gcg ggt cac ctc ctg acc gga aac gcc gcg agt				10416
Ala Ala Pro Glu Ala Ala Gly His Leu Leu Thr Gly Asn Ala Ala Ser				
3460	3465	3470		
gtc ctg gcc ggc cgg ctt tcc tac acg ttc ggc ctt gag ggg cct gcg				10464
Val Leu Ala Gly Arg Leu Ser Tyr Thr Phe Gly Leu Glu Gly Pro Ala				
3475	3480	3485		
gtg gcg gtg gat acc gcg tgt tcg tcg tcg ttg gtg gcg ttg cat ttg				10512
Val Ala Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu				
3490	3495	3500		
gcg tgc cag tcg ctg cgt tcg ggt gag tgt gat atg gcg ttg gca ggt				10560
Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Met Ala Leu Ala Gly				
3505	3510	3515	3520	
ggt gtg acg gtg atg tcg aca ccc ctg gct ttc ctc gag ttc tct cgt				10608
Gly Val Thr Val Met Ser Thr Pro Leu Ala Phe Leu Glu Phe Ser Arg				
3525	3530	3535		
cag cgc ggt ttg gcg cca gat ggt cgg tgc aag tcg ttt gcg gcc gct				10656
Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala				
3540	3545	3550		
gcg gat ggc acc ggg tgg ggt gag ggt gcc ggc ctg gtg ttg ctg gag				10704
Ala Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu				
3555	3560	3565		
cgg ttg tcg gat gct cgt cgg aat ggt cac cgg gtg ttg gcc gtg gtt				10752
Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Val				
3570	3575	3580		
cgc ggg tct gcg gtg aat cag gat ggt gcg tcg aat ggc ctg act gcg				10800
Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala				
3585	3590	3595	3600	
ccg aat ggt ccg tcg cag cag cgg gtg att cgg cag gcc ctc gcg aat				10848
Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn				
3605	3610	3615		
gcg ggg ctg tcg gcg tcc gat gtg gat gtc gtg gag gcg cac ggg acc				10896
Ala Gly Leu Ser Ala Ser Asp Val Asp Val Val Glu Ala His Gly Thr				

3620

3625

3630

ggt acc ggg ctc ggg gat ccg atc gag gcg cag gcg ctg atc gcg aca 10944  
 Gly Thr Gly Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr  
 3635 3640 3645

tat ggg cag gag cgg gat cct gag cgg gcc ctg tgg ctg ggg tcg atc 10992  
 Tyr Gly Gln Glu Arg Asp Pro Glu Arg Ala Leu Trp Leu Gly Ser Ile  
 3650 3655 3660

aag tcc aac atc ggc cac acg cag gcg gcg gcc ggt gtg gcg ggg gtc 11040  
 Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val  
 3665 3670 3675 3680

atc aag atg gtg cag gcc atg cgg cac ggg gag ttg cct gcg acg ttg 11088  
 Ile Lys Met Val Gln Ala Met Arg His Gly Glu Leu Pro Ala Thr Leu  
 3685 3690 3695

cac gtg gac aag ccc act cca cag gtg gac tgg tct gcc ggg gcc gtt 11136  
 His Val Asp Lys Pro Thr Pro Gln Val Asp Trp Ser Ala Gly Ala Val  
 3700 3705 3710

cgg ctc ctc acc ggg aac acg ccc tgg ccc gag agc ggc cgt cct cgt 11184  
 Arg Leu Leu Thr Gly Asn Thr Pro Trp Pro Glu Ser Gly Arg Pro Arg  
 3715 3720 3725

cga gcg ggg gtg tcg tcg ttc ggg atc agc ggc acc aac gca cac ctc 11232  
 Arg Ala Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu  
 3730 3735 3740

atc ctc gaa caa cca ccg tcg gaa cca gcg gag atc gac caa tcg gat 11280  
 Ile Leu Glu Gln Pro Pro Ser Glu Pro Ala Glu Ile Asp Gln Ser Asp  
 3745 3750 3755 3760

cgg cgg gtc act gcg cat cca gcg gtg atc ccg tgg atg ttg tcg gct 11328  
 Arg Arg Val Thr Ala His Pro Ala Val Ile Pro Trp Met Leu Ser Ala  
 3765 3770 3775

agg agt ctc gca gcg ctg cag gcc caa gcg gct gcg ctg cag gcc cgg 11376  
 Arg Ser Leu Ala Ala Leu Gln Ala Gln Ala Ala Leu Gln Ala Arg  
 3780 3785 3790

ctg gac cgg ggt cct ggc gct tct ccg ctg gat ttg ggg tat tca ctc 11424  
 Leu Asp Arg Gly Pro Gly Ala Ser Pro Leu Asp Leu Gly Tyr Ser Leu  
 3795 3800 3805

gcg acc act cgt tct gtg ctg gac gaa cgc gcc gtc gtg tgg ggt gcc 11472  
 Ala Thr Thr Arg Ser Val Leu Asp Glu Arg Ala Val Val Trp Gly Ala



3810	3815	3820	
gat cgg gag gca ctg ctg tcc agg ctg gca gcg ctc gcc gat ggc cgg			11520
Asp Arg Glu Ala Leu Leu Ser Arg Leu Ala Ala Leu Ala Asp Gly Arg			
3825	3830	3835	3840
acg gcg ccg ggg gtg ata acg ggc tct gcg aat tcc ggt ggc cgc atc			11568
Thr Ala Pro Gly Val Ile Thr Gly Ser Ala Asn Ser Gly Gly Arg Ile			
3845	3850		3855
gga ttc gtt ttt tcc ggt cag ggc agt cag tgg ctg ggg atg gga aag			11616
Gly Phe Val Phe Ser Gly Gln Gly Ser Gln Trp Leu Gly Met Gly Lys			
3860	3865		3870
gcg ttg tgc gcg gct ttc ccg gcg ttc gcg gac gcc ttc gag gaa gcc			11664
Ala Leu Cys Ala Ala Phe Pro Ala Phe Ala Asp Ala Phe Glu Glu Ala			
3875	3880		3885
tgc gac gcg cta agc gca cac ctg ggc gcg gac gtt cgg ggt gtg ctg			11712
Cys Asp Ala Leu Ser Ala His Leu Gly Ala Asp Val Arg Gly Val Leu			
3890	3895		3900
ttc ggt gct gat gag cag atg ctc gac ccg acg ctg tgg gcg cag tcg			11760
Phe Gly Ala Asp Glu Gln Met Leu Asp Arg Thr Leu Trp Ala Gln Ser			
3905	3910	3915	3920
ggg atc ttc gcg gtt caa gtc ggc ctc ctg gga ttg ctg agg tcg tgg			11808
Gly Ile Phe Ala Val Gln Val Gly Leu Leu Gly Leu Leu Arg Ser Trp			
3925	3930		3935
ggc gtg ccg ccg gcc gcg gtg ctg ggg cac tcg gtc ggc gag ttg gct			11856
Gly Val Arg Pro Ala Ala Val Leu Gly His Ser Val Gly Glu Leu Ala			
3940	3945		3950
gcg gcg cac gcg gct ggt gtg ttg tcc ttg ccg gac gct gca ccg ttg			11904
Ala Ala His Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Leu			
3955	3960		3965
gtt gcg gct ccg gcc cac ctg atg cag gca ttg ccc acc ggc ggc gca			11952
Val Ala Ala Arg Ala His Leu Met Gln Ala Leu Pro Thr Gly Gly Ala			
3970	3975		3980
atg ctc gcg gtc gcc acc agc gag gcg gcg gtc gga ccg ctg ctt tcc			12000
Met Leu Ala Val Ala Thr Ser Glu Ala Ala Val Gly Pro Leu Leu Ser			
3985	3990	3995	4000
ggg gtg tgc gat ccg gtc agc atc gct gcg atc aac ggc ccc gag tcg			12048
Gly Val Cys Asp Arg Val Ser Ile Ala Ala Ile Asn Gly Pro Glu Ser			

4005

4010

4015

gta gtg ctc tcc ggc gac cgc gat gtg ctc gtg gag ctc gca ggc gaa 12096  
 Val Val Leu Ser Gly Asp Arg Asp Val Leu Val Glu Leu Ala Gly Glu  
 4020 4025 4030

ttc gat gcc cga ggg ctt agg acc aaa tgg ttg cgg gtc tcc cat gct 12144  
 Phe Asp Ala Arg Gly Leu Arg Thr Lys Trp Leu Arg Val Ser His Ala  
 4035 4040 4045

ttc cac tcg cac cgg atg gaa ccg att ctg gac gag tac gcg gaa acc 12192  
 Phe His Ser His Arg Met Glu Pro Ile Leu Asp Glu Tyr Ala Glu Thr  
 4050 4055 4060

gcc agg tgc gtc gag ttc ggt gaa ccg gtg gtg ccg atc gtc tcc gcc 12240  
 Ala Arg Cys Val Glu Phe Gly Glu Pro Val Val Pro Ile Val Ser Ala  
 4065 4070 4075 4080

gcg acc ggt gcg ctg gac acc acc gga ctg atg tgc gcg gcc gac tac 12288  
 Ala Thr Gly Ala Leu Asp Thr Thr Gly Leu Met Cys Ala Ala Asp Tyr  
 4085 4090 4095

tgg acg cgc caa gtg cgt gat cct gtc cgc ttc gga gac ggt gtc cgg 12336  
 Trp Thr Arg Gln Val Arg Asp Pro Val Arg Phe Gly Asp Gly Val Arg  
 4100 4105 4110

gcg ctc gtc ggc caa ggc gtg gac acg atc gtc gag ttc ggc ccg gac 12384  
 Ala Leu Val Gly Gln Gly Val Asp Thr Ile Val Glu Phe Gly Pro Asp  
 4115 4120 4125

ggg gcg ttg tcg gcc ctg gtc gag cag tgc ttg gcc ggg tcc gac cag 12432  
 Gly Ala Leu Ser Ala Leu Val Glu Gln Cys Leu Ala Gly Ser Asp Gln  
 4130 4135 4140

gct ggg agg gtg gcg gcg atc ccg ctg atg cgc agg gac cgc gat gag 12480  
 Ala Gly Arg Val Ala Ala Ile Pro Leu Met Arg Arg Asp Arg Asp Glu  
 4145 4150 4155 4160

gtc gag acc gcg gtg gcg gcc ctg gcg cac gtg cac gtc cgc ggt ggt 12528  
 Val Glu Thr Ala Val Ala Ala Leu Ala His Val His Val Arg Gly Gly  
 4165 4170 4175

gcg gtg gac tgg tcg gct tgc ttc gcc ggc acc ggc gcc cgc acc gtc 12576  
 Ala Val Asp Trp Ser Ala Cys Phe Ala Gly Thr Gly Ala Arg Thr Val  
 4180 4185 4190

gag ttg ccc acc tac gcc ttc caa cgc cag cgg tac tgg ctg gcc ggg 12624  
 Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Ala Gly

4195	4200	4205	
caa gcg gac ggg cgc ggc ggc gat gtg gtt gcc gac ccg gtc gac gcg Gln Ala Asp Gly Arg Gly Gly Asp Val Val Ala Asp Pro Val Asp Ala 4210 4215 4220			12672
cgc ttc tgg gag ttg gtc gag cgc gcc gat ccg gaa ccg ttg gtg gat Arg Phe Trp Glu Leu Val Glu Arg Ala Asp Pro Glu Pro Leu Val Asp 4225 4230 4235 4240			12720
gaa ctc tgc atc gac cgg gac cag ccc ttc cgg gag gtg ctg ccc gtt Glu Leu Cys Ile Asp Arg Asp Gln Pro Phe Arg Glu Val Leu Pro Val 4245 4250 4255			12768
ctg gct tcc tgg cgc gag aaa caa cgc cag gag gcc ctc gcg gat tcc Leu Ala Ser Trp Arg Glu Lys Gln Arg Gln Glu Ala Leu Ala Asp Ser 4260 4265 4270			12816
tgg cgc tac cag gtg cgc tgg agg tcc gtc gag gtg ccg tcc gca gcc Trp Arg Tyr Gln Val Arg Trp Arg Ser Val Glu Val Pro Ser Ala Ala 4275 4280 4285			12864
gcc ctc cgg ggc gtg tgg ctg gtg gtg ctt cca gct gac gtg ccc cga Ala Leu Arg Gly Val Trp Leu Val Val Leu Pro Ala Asp Val Pro Arg 4290 4295 4300			12912
gat caa ccg gcg gtc gtc atc gac gcg ctg atc gcg cgc ggc gcc gag Asp Gln Pro Ala Val Val Ile Asp Ala Leu Ile Ala Arg Gly Ala Glu 4305 4310 4315 4320			12960
gtc gcg gtc ctg gaa ttg acc gag cag gac ctc caa cgc agt gcg ctt Val Ala Val Leu Glu Leu Thr Glu Gln Asp Leu Gln Arg Ser Ala Leu 4325 4330 4335			13008
gtg gac aag gtg cgc gcc gtc att gcg gac cgc acc gag gtg acg ggt Val Asp Lys Val Arg Ala Val Ile Ala Asp Arg Thr Glu Val Thr Gly 4340 4345 4350			13056
gtg ttg tct ctg ttg gcg atg gac ggc atg ccc tgc gcg gcg cat ccg Val Leu Ser Leu Leu Ala Met Asp Gly Met Pro Cys Ala Ala His Pro 4355 4360 4365			13104
cac ctg tcc cgt ggt gtc gcc gct acc gtg atc ctg acg cag gtg ttg His Leu Ser Arg Gly Val Ala Ala Thr Val Ile Leu Thr Gln Val Leu 4370 4375 4380			13152
ggc gat gcg ggt gtt tcc gcc ccg ctg tgg ctg gcc acg acc ggt ggc Gly Asp Ala Gly Val Ser Ala Pro Leu Trp Leu Ala Thr Thr Gly Gly			13200

4385

4390

4395

4400

gtc gag gcc ggg acc gag gac ggt ccg gcc gat ccg gac cac ggc ttg 13248  
 Val Glu Ala Gly Thr Glu Asp Gly Pro Ala Asp Pro Asp His Gly Leu  
 4405 4410 4415

atc tgg ggg ctc ggc agg gtc gtc ggc ctt gaa cat ccg cag tgg tgg 13296  
 Ile Trp Gly Leu Gly Arg Val Val Gly Leu Glu His Pro Gln Trp Trp  
 4420 4425 4430

ggt ggc ctg atc gac ctt ccg gag aca ctg gac gag acg tcc cgg aac 13344  
 Gly Gly Leu Ile Asp Leu Pro Glu Thr Leu Asp Glu Thr Ser Arg Asn  
 4435 4440 4445

ggg ttg gtg gcc gca ctc gcc ggg acg gcg gcc gaa gat cag ctc gcc 13392  
 Gly Leu Val Ala Ala Leu Ala Gly Thr Ala Ala Glu Asp Gln Leu Ala  
 4450 4455 4460

gtg cgt tca tcc ggg ttg ttc gtt cgc aga gtg gtg cgc gca gcg cgg 13440  
 Val Arg Ser Ser Gly Leu Phe Val Arg Arg Val Val Arg Ala Ala Arg  
 4465 4470 4475 4480

aac ccc cgg tca gag aca tgg cgt agc cgg gga acg gtc ctc atc acg 13488  
 Asn Pro Arg Ser Glu Thr Trp Arg Ser Arg Gly Thr Val Leu Ile Thr  
 4485 4490 4495

ggc gga aca ggc gcg ctc ggt gcc gag gtc gca cga tgg ctg gcc cgg 13536  
 Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg Trp Leu Ala Arg  
 4500 4505 4510

cgg gga gct gag cac ctg gtg ttg atc agt cgc cgc ggc ccg gaa gct 13584  
 Arg Gly Ala Glu His Leu Val Leu Ile Ser Arg Arg Gly Pro Glu Ala  
 4515 4520 4525

ccc ggc gca gcg gac cta ggg gcc gag ctg act gaa ctc ggc gtg aaa 13632  
 Pro Gly Ala Ala Asp Leu Gly Ala Glu Leu Thr Glu Leu Gly Val Lys  
 4530 4535 4540

gtc aca gtc ttg gcc tgc gat gtg acg gac cgc gac gag ctg gcg gcg 13680  
 Val Thr Val Leu Ala Cys Asp Val Thr Asp Arg Asp Glu Leu Ala Ala  
 4545 4550 4555 4560

gtg ctg gcg gcc gtt ccc acg gag tat ccg ctg tcg gcg gtc gtg cac 13728  
 Val Leu Ala Ala Val Pro Thr Glu Tyr Pro Leu Ser Ala Val Val His  
 4565 4570 4575

acc gcc ggc gtc ggg acg cct gcg aac ctg gcc gag acg acc ttg gcg 13776  
 Thr Ala Gly Val Gly Thr Pro Ala Asn Leu Ala Glu Thr Thr Leu Ala

4580	4585	4590	
cag ttc gcc gac gtg ttg tcg gcc aag gtc gtc ggc gcg gcg aac ctg			13824
Gln Phe Ala Asp Val Leu Ser Ala Lys Val Val Gly Ala Ala Asn Leu			
4595	4600	4605	
gac cgg ctg ctt ggc ggg caa ccg ttg gac gcc ttc gtg ctg ttc tcc			13872
Asp Arg Leu Leu Gly Gly Gln Pro Leu Asp Ala Phe Val Leu Phe Ser			
4610	4615	4620	
tcg atc tcg gga gtt tgg gga gcc ggc ggc caa gga gcc tat tcg gcc			13920
Ser Ile Ser Gly Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ser Ala			
4625	4630	4635	4640
gcc aat gcg tat ctc gat gcc ctt gcc gag cgc cga cgg gct tgc ggg			13968
Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Arg Arg Arg Ala Cys Gly			
4645	4650	4655	
cgg ccg gcg acg tgc atc gcc tgg ggt ccg tgg gcg ggt gcg ggc atg			14016
Arg Pro Ala Thr Cys Ile Ala Trp Gly Pro Trp Ala Gly Ala Gly Met			
4660	4665	4670	
gcc gtt cag gaa ggt aac gag gcg cat ctc cgc cga agg ggc ctg gta			14064
Ala Val Gln Glu Gly Asn Glu Ala His Leu Arg Arg Arg Gly Leu Val			
4675	4680	4685	
ccg atg gaa ccg cag tcg gcc ctc ttc gcg ctg caa cag gcc ctg tcc			14112
Pro Met Glu Pro Gln Ser Ala Leu Phe Ala Leu Gln Gln Ala Leu Ser			
4690	4695	4700	
caa cga gaa acc gcc atc acc gtc gca gat gtg gac tgg gag cga ttc			14160
Gln Arg Glu Thr Ala Ile Thr Val Ala Asp Val Asp Trp Glu Arg Phe			
4705	4710	4715	4720
gcc gcc tct ttc acc gcg gcc cgc ccg cga cca ctg ttg gaa gag atc			14208
Ala Ala Ser Phe Thr Ala Ala Arg Pro Arg Pro Leu Leu Glu Glu Ile			
4725	4730	4735	
gtg gat cta cgg ccc gac acc gag acc gag gag aag cac ggt gcc ggc			14256
Val Asp Leu Arg Pro Asp Thr Glu Thr Glu Glu Lys His Gly Ala Gly			
4740	4745	4750	
gag ctg ggg cag cag ctg gcc gca ctg ccg ccc gct gag cgc gga cac			14304
Glu Leu Gly Gln Gln Leu Ala Ala Leu Pro Pro Ala Glu Arg Gly His			
4755	4760	4765	
ctg ctg ctg gag gtg gtg ctg gcg gaa acc gcc agc acc ctg ggg cac			14352
Leu Leu Leu Glu Val Val Leu Ala Glu Thr Ala Ser Thr Leu Gly His			

4770

4775

4780

gat tcg gcg gag gct gtg caa ccc gat cgg acc ttc gcc gaa ctg ggc 14400  
 Asp Ser Ala Glu Ala Val Gln Pro Asp Arg Thr Phe Ala Glu Leu Gly  
 4785 4790 4795 4800

ttc gat tcg ctg acc gcg gta gag ctg cgc aac agg ttg aac gcg gtg 14448  
 Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Asn Ala Val  
 4805 4810 4815

acc ggg ctt cgc ctg ccg ccg acg ctg gtt ttc gac cac ccg acg ccg 14496  
 Thr Gly Leu Arg Leu Pro Pro Thr Leu Val Phe Asp His Pro Thr Pro  
 4820 4825 4830

ctg gcg ttg tcc gaa cag ttg gtt ccg gcc ctg gtc gcg gag ccg gac 14544  
 Leu Ala Leu Ser Glu Gln Leu Val Pro Ala Leu Val Ala Glu Pro Asp  
 4835 4840 4845

aac ggc atc gaa tcg ctg ctc gcc gag ctc gac agg ctg gat acc acg 14592  
 Asn Gly Ile Glu Ser Leu Leu Ala Glu Leu Asp Arg Leu Asp Thr Thr  
 4850 4855 4860

ttg gcg caa ggg cct tcg atc cca ctg gaa gac cag gcc aag gtg gcg 14640  
 Leu Ala Gln Gly Pro Ser Ile Pro Leu Glu Asp Gln Ala Lys Val Ala  
 4865 4870 4875 4880

gag cgc ttg cac gca ctc ctc gcc aag tgg gac ggg gcg cgt gac ggc 14688  
 Glu Arg Leu His Ala Leu Leu Ala Lys Trp Asp Gly Ala Arg Asp Gly  
 4885 4890 4895

acg gcc aga gcg acg tca ccc caa tcg ctg acg gcg gcc acg gac gac 14736  
 Thr Ala Arg Ala Thr Ser Pro Gln Ser Leu Thr Ala Ala Thr Asp Asp  
 4900 4905 4910

gaa atc ttc gac ctc atc gac cgg aag ttc cgg cgc tga 14775  
 Glu Ile Phe Asp Leu Ile Asp Arg Lys Phe Arg Arg  
 4915 4920

&lt;210&gt; 48

&lt;211&gt; 4924

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 48

Met Ala Asn Glu Glu Lys Leu Phe Gly Tyr Leu Lys Lys Val Thr Ala  
 1 5 10 15

Asp Leu His Gln Thr Arg Gln Arg Leu Leu Ala Ala Glu Ser Arg Ser  
20 25 30

Gln Glu Pro Ile Ala Ile Val Ser Ala Ser Cys Arg Leu Pro Gly Gly  
35 40 45

Val Asp Ser Pro Glu Ala Leu Trp Gln Leu Val Arg Thr Gly Thr Asp  
50 55 60

Ala Ile Ser Glu Phe Pro Ala Asp Arg Gly Trp Asp Leu Gly Arg Leu  
65 70 75 80

Tyr Asp Pro Asp Pro Asn His Gln Gly Thr Ser Tyr Thr Arg Ala Gly  
85 90 95

Gly Phe Leu Ala Gly Ala Gly Asp Phe Asp Pro Ala Met Phe Gly Ile  
100 105 110

Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
115 120 125

Glu Leu Ser Trp Glu Ala Leu Glu Arg Ala Gly Ile Asp Pro Thr Ser  
130 135 140

Leu Arg Gly Ser Lys Thr Gly Val Phe Gly Gly Val Thr Pro Gln Glu  
145 150 155 160

Tyr Gly Pro Ser Leu Gln Glu Met Ser Arg Asn Ala Gly Gly Phe Gly  
165 170 175

Leu Thr Gly Arg Met Val Ser Val Ala Ser Gly Arg Val Ala Tyr Ser  
180 185 190

Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser  
195 200 205

Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Ser Gly Glu  
210 215 220

Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Ala  
225 230 235 240

Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg  
245 250 255

Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly  
260 265 270

Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly  
275 280 285

His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly  
290 295 300

Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val  
305 310 315 320

Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser Val Ser Asp Val Asp  
325 330 335

Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu  
340 345 350

Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly Arg Glu Lys Asp Arg  
355 360 365

Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Thr Gln Ala  
370 375 380

Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Leu Ala Met Arg His  
385 390 395 400

Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu Pro Thr Ser Ala Val  
405 410 415

Asp Trp Ser Ala Gly Ser Val Arg Leu Leu Thr Glu Asn Thr Pro Trp  
420 425 430

Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val Ser Ser Phe Gly Ile  
435 440 445

Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ser Pro Val Glu Gln  
450 455 460

Gly Glu Pro Ala Gly Pro Val Glu Gly Glu Arg Glu Pro Asp Val Ala  
465 470 475 480

Val Pro Val Val Pro Trp Val Leu Ser Gly Lys Thr Pro Glu Ala Ala  
485 490 495

Arg Ala Gln Ala Glu Arg Val His Ser His Ile Glu Asp Arg Pro Gly  
500 505 510

Leu Ser Pro Val Asp Val Ala Tyr Ser Leu Gly Met Thr Arg Ala Ala  
515 520 525



Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Ala Ala Leu Leu  
530 535 540

Thr Gly Leu Arg Ala Phe Ala Asp Gly Cys Asp Ala Pro Glu Val Val  
545 550 555 560

Ser Gly Ser Val Gly Leu Gly Gly Arg Val Gly Phe Val Phe Ser Gly  
565 570 575

Gln Gly Gly Gln Trp Pro Gly Met Gly Arg Gly Leu Tyr Ser Val Phe  
580 585 590

Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp Ala  
595 600 605

His Leu Gly Gln Glu Leu Arg Val Arg Asp Val Val Phe Gly Ser Gln  
610 615 620

Ala Trp Leu Leu Asp Arg Thr Val Trp Ala Gln Ser Gly Leu Phe Ala  
625 630 635 640

Leu Gln Ile Gly Leu Leu Arg Leu Leu Gly Ser Trp Gly Val Arg Pro  
645 650 655

Asp Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Val His Ala  
660 665 670

Ala Gly Val Leu Ser Leu Ser Glu Ala Ala Arg Leu Val Ala Gly Arg  
675 680 685

Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala Met Leu Ala Val  
690 695 700

Ala Thr Gly Glu Phe Gln Val Asp Pro Leu Leu Asp Gly Val Arg Asp  
705 710 715 720

Arg Ile Gly Ile Ala Ala Val Asn Gly Pro Glu Ser Val Val Leu Ser  
725 730 735

Gly Asp Arg Glu Leu Leu Thr Glu Ile Ala Asp Arg Leu His Asp Gln  
740 745 750

Gly Cys Arg Thr Arg Trp Leu Arg Val Ser His Ala Phe His Ser Pro  
755 760 765

His Met Glu Pro Met Leu Glu Glu Phe Ala Gln Ile Ser Arg Gly Arg  
770 775 780

Glu Tyr His Ala Phe Glu Leu Pro Ile Ile Ser Thr Leu Ile Gly Glu  
 785 790 795 800

Leu Asp Gly Gly Arg Val Met Gly Thr Pro Glu Tyr Trp Val Arg Gln  
 805 810 815

Val Arg Glu Pro Val Arg Phe Ala Glu Gly Val Gln Ala Leu Val Gly  
 820 825 830

Gln Gly Val Gly Thr Ile Val Glu Leu Gly Pro Asp Gly Ala Leu Ser  
 835 840 845

Thr Leu Val Glu Glu Cys Val Ala Glu Ser Gly Arg Val Ala Gly Ile  
 850 855 860

Pro Leu Met Arg Lys Asp Arg Asp Glu Ala Arg Thr Val Leu Ala Ala  
 865 870 875 880

Leu Ala Gln Ile His Thr Arg Gly Gly Glu Val Asp Trp Arg Ser Phe  
 885 890 895

Phe Ala Gly Thr Gly Ala Lys Gln Val Asp Leu Pro Thr Tyr Ala Phe  
 900 905 910

Gln Arg Gln Arg Tyr Trp Leu Ala Ser Thr Gly Arg Ala Gly Asp Val  
 915 920 925

Thr Ala Ala Gly Leu Ala Glu Ala Asp His Pro Leu Leu Gly Ala Val  
 930 935 940

Val Ala Leu Ala Asp Gly Glu Gly Val Val Leu Thr Gly Arg Leu Thr  
 945 950 955 960

Ala Gly Ser His Pro Trp Leu Ser Asp His Arg Val Leu Gly Glu Ile  
 965 970 975

Val Val Pro Gly Thr Ala Ile Val Glu Leu Val Trp His Val Gly Glu  
 980 985 990

Arg Leu Gly Cys Gly Arg Val Glu Glu Leu Ala Leu Glu Ala Pro Leu  
 995 1000 1005

Ile Leu Pro Asp His Gly Ala Val Gln Val Gln Val Leu Val Gly Pro  
 1010 1015 1020

Pro Gly Glu Ser Gly Ala Arg Ser Val Ala Leu Tyr Ser Cys Pro Gly  
 1025 1030 1035 1040

Glu Ala Ile Glu Pro Glu Trp Lys Lys His Ala Thr Gly Val Leu Leu  
1045 1050 1055

Pro Pro Val Ala Ala Glu Asn His Glu Leu Thr Ala Trp Pro Pro Glu  
1060 1065 1070

Asn Ala Thr Glu Ile Asp Ala Asp Gly Val Tyr Ala Phe Leu Glu Gly  
1075 1080 1085

His Gly Phe Ala Tyr Gly Pro Ala Phe Arg Cys Leu Arg Gly Ala Trp  
1090 1095 1100

Arg Arg Gly Gly Glu Val Phe Ala Glu Val Ala Leu Pro Asp Asp Met  
1105 1110 1115 1120

Gln Ala Gly Val Asp Arg Phe Gly Val His Pro Ala Leu Leu Asp Ala  
1125 1130 1135

Val Leu His Ala Ala Ala Glu Thr Ser Val Val Gln Ser Glu Ala  
1140 1145 1150

Arg Val Pro Phe Ser Trp Arg Gly Val Glu Leu Arg Ala Thr Glu Ser  
1155 1160 1165

Ala Val Val Arg Ala Arg Leu Ser Leu Thr Ser Asp Asp Glu Leu Ser  
1170 1175 1180

Leu Val Ala Val Asp Pro Ala Gly Arg Phe Val Ala Thr Val Asp Ser  
1185 1190 1195 1200

Leu Val Thr Arg Pro Ile Ser Arg Gln Gln Val Arg Ser Gly Ala Ile  
1205 1210 1215

Gly Asp Cys Leu Phe Glu Val Glu Trp His Arg Lys Ala Leu Leu Gly  
1220 1225 1230

Thr Thr Ala Gly Asp Asp Leu Ala Ile Val Gly Asp Gly Pro Ser Trp  
1235 1240 1245

Pro Glu Ser Val Arg Ala Thr Ala Arg Phe Ala Thr Leu Asp Glu Phe  
1250 1255 1260

Arg Ala Ala Val Asp Ser Asp Val Pro Ala Pro Gly Ser Val Leu Val  
1265 1270 1275 1280

Ala Ala Met Ser Ala Glu Glu Val Glu Gly Gly Ser Leu Pro Ser Arg  
1285 1290 1295

Ala Gln Glu Ser Thr Ser Asp Leu Leu Ala Leu Val Gln Ser Trp Leu  
1300 1305 1310

Ala Asp Glu Arg Phe Ala Glu Ser Gln Leu Val Val Val Thr Arg Ala  
1315 1320 1325

Ala Val Ser Ala Asp Ser Asp Ser Asp Val Ala Asp Leu Val Gly Ala  
1330 1335 1340

Ser Ser Trp Gly Leu Leu Ser Ser Ala Gln Ser Glu Asn Pro Gly Arg  
1345 1350 1355 1360

Phe Val Leu Val Asp Val Asp Gly Thr Pro Glu Ser Trp Gln Ala Leu  
1365 1370 1375

Pro Ala Ala Val Arg Ala Gly Glu Pro Gln Leu Ala Leu Arg Arg Gly  
1380 1385 1390

Ala Ala Leu Val Pro Arg Leu Ala Arg Leu Thr Val Arg Glu Glu Gly  
1395 1400 1405

Ser Ser Pro Gln Leu Asp Thr Asp Gly Thr Val Leu Ile Thr Gly Gly  
1410 1415 1420

Thr Gly Ala Leu Gly Gly Val Val Ala Arg His Leu Val Glu Glu His  
1425 1430 1435 1440

Gly Ile Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp Asn Ala Pro  
1445 1450 1455

Gly Val His Glu Leu Val Asp Glu Leu Ala Arg Ala Gly Ala Val Val  
1460 1465 1470

Leu Val Val Ala Cys Asp Val Ala Asp Arg Thr Asp Leu Glu His Val  
1475 1480 1485

Leu Ala Ala Ile Pro Val Asp Trp Pro Leu Arg Gly Ile Val His Thr  
1490 1495 1500

Ala Gly Val Leu Ala Asp Gly Val Ile Gly Ser Leu Ser Ala Ala Asp  
1505 1510 1515 1520

Val Gly Thr Val Phe Ala Pro Lys Val Thr Gly Ala Trp His Leu His  
1525 1530 1535

Glu Leu Thr Arg Asp Leu Asp Leu Ser Phe Phe Val Leu Phe Ser Ser  
1540 1545 1550

Phe Ser Gly Ile Ala Gly Ala Ala Gly Gln Ala Asn Tyr Ala Ala Ala  
1555 1560 1565

Asn Thr Phe Leu Asp Ala Leu Ala Arg Tyr Arg Arg Ala Arg Gly Leu  
1570 1575 1580

Pro Gly Leu Ser Leu Ala Trp Gly Leu Trp Ala Gln Pro Ser Gly Met  
585 1590 1595 1600

Thr Ser Gly Leu Asp Ala Ala Ser Val Glu Arg Leu Ala Arg Thr Gly  
1605 1610 1615

Ile Ala Glu Leu Ser Thr Glu Asp Gly Leu Arg Leu Phe Asp Ala Ala  
1620 1625 1630

Phe Ala Lys Asp Arg Ala Cys Val Val Ala Ala Arg Leu Asp Arg Ala  
1635 1640 1645

Leu Leu Val Gly Asn Gly Arg Ser His Ala Ile Pro Ala Leu Leu Ser  
1650 1655 1660

Ala Leu Val Pro Val Arg Gly Gly Val Ala Arg Lys Thr Ala Asn Ser  
665 1670 1675 1680

Gln Ala Ala Asp Glu Asp Ala Leu Leu Gly Leu Val Arg Glu His Val  
1685 1690 1695

Ser Ala Val Leu Gly Tyr Ser Gly Ala Val Glu Val Gly Gly Asp Arg  
1700 1705 1710

Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg  
1715 1720 1725

Arg Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val  
1730 1735 1740

Phe Asp Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu  
745 1750 1755 1760

Leu Ala Gly Glu Val Ala Ser Thr Ser Thr Pro Val Thr Arg Ala Ala  
1765 1770 1775

Ser Ala Glu Glu Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe  
1780 1785 1790

Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly  
1795 1800 1805

Gly Val Asp Ala Val Ala Gly Phe Pro Asp Asp Arg Gly Trp Asp Leu  
 1810 1815 1820

Ala Ala Leu Tyr Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr Val  
 825 1830 1835 1840

Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Asp Met  
 1845 1850 1855

Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg  
 1860 1865 1870

Leu Leu Leu Glu Val Ala Trp Glu Thr Leu Glu Arg Ala Gly Ile Asp  
 1875 1880 1885

Pro Phe Ser Leu His Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met  
 1890 1895 1900

His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly Phe  
 905 1910 1915 1920

Glu Gly His Leu Gly Thr Gly Asn Ala Gly Ser Val Leu Ser Gly Arg  
 1925 1930 1935

Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr  
 1940 1945 1950

Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala Leu  
 1955 1960 1965

Arg Ala Gly Glu Cys Glu Phe Ala Leu Ala Gly Gly Val Thr Val Met  
 1970 1975 1980

Thr Pro Thr Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala  
 1990 1995 2000

Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr Gly  
 2005 2010 2015

Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala  
 2020 2025 2030

Arg Arg Asn Gly His Glu Val Leu Ala Val Val Arg Gly Ser Ala Val  
 2035 2040 2045

Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser  
 2050 2055 2060

Gln Gln Arg Val Ile Thr Gln Ala Leu Thr Ser Ala Gly Leu Ser Val  
065 2070 2075 2080

Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly  
2085 2090 2095

Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg  
2100 2105 2110

Asp Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly  
2115 2120 2125

His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met  
2130 2135 2140

Ala Met Arg Gln Gly Glu Leu Pro Arg Thr Leu His Val Asp Glu Pro  
145 2150 2155 2160

Ser Ala Gln Val Asp Trp Ser Ala Gly Thr Val Gln Leu Leu Thr Glu  
2165 2170 2175

Asn Thr Pro Trp Pro Asp Ser Gly Arg Leu Arg Arg Ala Gly Val Ser  
2180 2185 2190

Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu Ile Leu Glu Gln Pro  
2195 2200 2205

Pro Arg Glu Ser Gln Arg Ser Thr Glu Pro Asp Ser Gly Ser Val Arg  
2210 2215 2220

Asp Phe Pro Val Val Pro Trp Met Val Ser Gly Lys Thr Pro Glu Ala  
225 2230 2235 2240

Leu Ser Ala Gln Ala Asp Ala Leu Met Ser Tyr Leu Ser Asn Arg Val  
2245 2250 2255

Asp Ala Ser Pro Arg Asp Ile Gly Tyr Ser Leu Ala Val Thr Arg Pro  
2260 2265 2270

Ala Leu Asp His Arg Ala Val Val Leu Gly Ala Asp Arg Ala Ala Leu  
2275 2280 2285

Leu Pro Gly Leu Lys Ala Leu Ala Val Ser Asn Asp Ala Ala Glu Val  
2290 2295 2300

Ile Thr Gly Thr Arg Ala Ala Gly Pro Val Gly Phe Val Phe Ser Gly  
305 2310 2315 2320

Gln Gly Gly Gln Thr Pro Gly Met Gly Ser Gly Leu His Ser Ala Phe  
2325 2330 2335

Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Cys Glu Leu Asp Ala  
2340 2345 2350

His Leu Gly Gln Met Ala Arg Leu Arg Asp Val Leu Ser Gly Ser Asp  
2355 2360 2365

Thr Gln Leu Leu Asp Gln Thr Leu Trp Ala Gln Pro Gly Leu Phe Ala  
2370 2375 2380

Leu Gln Val Gly Leu Trp Glu Leu Leu Gly Ser Trp Gly Val Arg Pro  
2385 2390 2395 2400

Ala Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe Ala  
2405 2410 2415

Ala Gly Val Leu Ser Leu Arg Asp Ala Ala Arg Leu Val Ala Gly Arg  
2420 2425 2430

Ala Arg Leu Met Gln Ala Leu Pro Thr Gly Gly Ala Met Leu Ala Ala  
2435 2440 2445

Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala Asp Cys Gly Asp  
2450 2455 2460

Arg Val Gly Ile Ala Ala Val Asn Ala Pro Gly Ser Val Val Leu Ser  
2465 2470 2475 2480

Gly Asp Arg Asp Val Leu Asp Asp Ile Ala Gly Arg Leu Asp Gly Gln  
2485 2490 2495

Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser His  
2500 2505 2510

Arg Met Asp Pro Met Leu Ala Glu Phe Thr Glu Ile Ala Arg Ser Val  
2515 2520 2525

Asp Tyr Arg Ser Ser Gly Leu Pro Ile Val Ser Thr Leu Thr Gly Glu  
2530 2535 2540

Leu Asp Glu Val Gly Met Pro Ala Thr Pro Glu Tyr Trp Val Arg Gln  
2545 2550 2555 2560

Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Ala Ala Leu Ala Ala  
2565 2570 2575



His Gly Val Ser Thr Val Val Glu Val Gly Pro Asp Gly Val Leu Ser  
2580 2585 2590

Ala Leu Val Gln Glu Cys Ala Ala Gly Ser Asp Gln Gly Gly Arg Val  
2595 2600 2605

Ala Ala Val Pro Leu Met Arg Ser Asn Arg Asp Glu Ala His Thr Val  
2610 2615 2620

Thr Thr Ala Leu Ala Gln Ile His Val Arg Gly Ala Glu Val Asp Trp  
625 2630 2635 2640

Arg Ser Phe Phe Ala Gly Thr Gly Ala Lys Gln Val Glu Leu Pro Thr  
2645 2650 2655

Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Asp Ser Pro Ser Glu Pro  
2660 2665 2670

Val Gly Gln Ser Ala Asp Pro Ala Arg Gln Ser Gly Phe Trp Glu Leu  
2675 2680 2685

Val Glu Gln Glu Asp Val Ser Ala Leu Ser Ala Ala Leu His Ile Thr  
2690 2695 2700

Gly Asp His Asp Val Gln Ala Ser Leu Glu Ser Val Val Pro Val Leu  
705 2710 2715 2720

Ser Ser Trp His Arg Arg Ile Arg Asn Glu Ser Leu Val His Gln Trp  
2725 2730 2735

Arg Tyr Arg Ile Ser Trp His Glu Arg Ala Asp Leu Pro Asp Pro Ser  
2740 2745 2750

Leu Ser Gly Thr Trp Leu Val Val Val Pro Glu Gly Trp Ser Ala Ser  
2755 2760 2765

Arg Gln Val Leu Arg Phe Asn Glu Met Phe Glu Glu Arg Gly Cys Pro  
2770 2775 2780

Ala Val Leu Phe Glu Leu Ala Gly His Asp Glu Glu Ala Leu Ala Gln  
785 2790 2795 2800

Arg Phe Arg Ser Leu Pro Val Ala Ser Gly Gly Ile Ser Gly Val Leu  
2805 2810 2815

Ser Leu Leu Ala Leu Asp Glu Ser Pro Ser Ser Pro Asn Ala Ala Leu  
2820 2825 2830

Pro Asn Gly Ala D Asn Ser Leu Val Leu Leu Arg A Leu Arg Ala  
 2835 2840 2845

Ala Asp Val Ser Ala Pro Leu Trp Leu Ala Thr Cys Gly Gly Val Ala  
 2850 2855 2860

Val Gly Asp Val Pro Val Asn Pro Gly Gln Ala Leu Val Trp Gly Leu  
 865 2870 2875 2880

Gly Arg Val Val Gly Leu Glu His Pro Ala Trp Trp Gly Gly Leu Val  
 2885 2890 2895

Asp Val Pro Cys Leu Leu Asp Glu Asp Ala Arg Glu Arg Leu Ser Val  
 2900 2905 2910

Val Leu Ala Gly Leu Gly Glu Asp Glu Ile Ala Val Arg Pro Gly Gly  
 2915 2920 2925

Val Phe Val Arg Arg Leu Glu Arg Ala Gly Ala Ala Ser Gly Ala Gly  
 2930 2935 2940

Ser Val Trp Arg Pro Arg Gly Thr Val Leu Val Thr Gly Gly Thr Gly  
 945 2950 2955 2960

Gly Leu Gly Ala His Val Ala Arg Trp Leu Ala Gly Ala Gly Ala Glu  
 2965 2970 2975

His Val Val Leu Thr Ser Arg Arg Gly Ala Ala Ala Pro Gly Ala Gly  
 2980 2985 2990

Asp Leu Arg Ala Glu Leu Glu Ala Leu Gly Ala Arg Val Ser Ile Thr  
 2995 3000 3005

Ala Cys Asp Val Ala Asp Arg Asp Ala Leu Ala Glu Val Leu Ala Thr  
 3010 3015 3020

Ile Pro Asp Asp Cys Pro Leu Thr Ala Val Met His Ala Ala Gly Val  
 025 3030 3035 3040

Val Glu Val Gly Asp Val Ala Ser Met Cys Leu Thr Asp Phe Val Gly  
 3045 3050 3055

Val Leu Ser Ala Lys Ala Gly Gly Ala Ala Asn Leu Asp Glu Leu Leu  
 3060 3065 3070

Ala Asp Val Glu Leu Asp Ala Phe Val Leu Phe Ser Ser Val Ser Gly  
 3075 3080 3085

Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ala Ala Ala Asn Ala Tyr  
3090 3095 3100

Leu Asp Ala Leu Ala Gln Gln Arg Arg Ala Arg Gly Leu Val Gly Thr  
105 3110 3115 3120

Ala Val Ala Trp Gly Pro Trp Ala Gly Asp Gly Met Ala Ala Gly Glu  
3125 3130 3135

Gly Gly Ala Gln Leu Arg Arg Ala Gly Leu Val Pro Met Ala Ala Asp  
3140 3145 3150

Arg Ala Leu Leu Ala Leu Gln Gly Ala Leu Asp Arg Asp Glu Thr Ser  
3155 3160 3165

Leu Val Val Ala Asp Met Ala Trp Glu Arg Phe Ala Pro Val Phe Ala  
3170 3175 3180

Met Ser Arg Arg Arg Pro Leu Leu Asp Glu Leu Pro Glu Ala Gln Gln  
185 3190 3195 3200

Ala Leu Ala Asp Ala Glu Asn Thr Thr Asp Ala Ala Asp Ser Ala Val  
3205 3210 3215

Pro Leu Pro Arg Leu Ala Gly Met Ala Ala Ala Glu Arg Arg Arg Ala  
3220 3225 3230

Met Leu Asp Leu Val Leu Ala Glu Ala Ser Ile Val Leu Gly His Asn  
3235 3240 3245

Gly Ser Asp Pro Val Gly Pro Asp Arg Ala Phe Gln Glu Leu Gly Phe  
3250 3255 3260

Asp Ser Leu Met Ala Val Glu Leu Arg Asn Arg Leu Gly Glu Ala Thr  
265 3270 3275 3280

Gly Leu Ser Leu Pro Ala Thr Leu Ile Phe Asp Tyr Pro Ser Pro Ser  
3285 3290 3295

Ala Leu Ala Glu Gln Leu Val Gly Glu Leu Val Gly Ala Gln Pro Ala  
3300 3305 3310

Thr Thr Val Val Ala Gly Ala Asp Pro Val Asp Asp Pro Val Val Val  
3315 3320 3325

Val Ala Met Gly Cys Arg Tyr Pro Gly Asp Val Cys Ser Pro Glu Glu  
3330 3335 3340

Leu Trp Gln Leu Val Ser Ala Gly Arg Asp Ala Val Ser Thr Phe Pro  
 345 3350 3355 3360

Val Asp Arg Gly Trp Asp Cys Asn Thr Leu Phe Asp Pro Asp Pro Asp  
 3365 3370 3375

Arg Ala Gly Ser Thr Tyr Val Arg Glu Gly Ala Phe Leu Thr Gly Ala  
 3380 3385 3390

Asp Arg Phe Asp Ala Gly Phe Phe Gly Ile Ser Pro Arg Glu Ala Arg  
 3395 3400 3405

Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Val Ala Trp Glu Val  
 3410 3415 3420

Phe Glu Arg Ala Gly Ile Ala Pro Leu Ser Leu Arg Gly Ser Arg Thr  
 425 3430 3435 3440

Gly Val Phe Ala Gly Thr Asn Gly Gln Asp His Gly Ala Lys Val Ala  
 3445 3450 3455

Ala Ala Pro Glu Ala Ala Gly His Leu Leu Thr Gly Asn Ala Ala Ser  
 3460 3465 3470

Val Leu Ala Gly Arg Leu Ser Tyr Thr Phe Gly Leu Glu Gly Pro Ala  
 3475 3480 3485

Val Ala Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu  
 3490 3495 3500

Ala Cys Gln Ser Leu Arg Ser Gly Glu Cys Asp Met Ala Leu Ala Gly  
 505 3510 3515 3520

Gly Val Thr Val Met Ser Thr Pro Leu Ala Phe Leu Glu Phe Ser Arg  
 3525 3530 3535

Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala  
 3540 3545 3550

Ala Asp Gly Thr Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu  
 3555 3560 3565

Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Val  
 3570 3575 3580

Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala  
 585 3590 3595 3600

Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn  
3605 3610 3615

Ala Gly Leu Ser Ala Ser Asp Val Asp Val Val Glu Ala His Gly Thr  
3620 3625 3630

Gly Thr Gly Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr  
3635 3640 3645

Tyr Gly Gln Glu Arg Asp Pro Glu Arg Ala Leu Trp Leu Gly Ser Ile  
3650 3655 3660

Lys Ser Asn Ile Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val  
3665 3670 3675 3680

Ile Lys Met Val Gln Ala Met Arg His Gly Glu Leu Pro Ala Thr Leu  
3685 3690 3695

His Val Asp Lys Pro Thr Pro Gln Val Asp Trp Ser Ala Gly Ala Val  
3700 3705 3710

Arg Leu Leu Thr Gly Asn Thr Pro Trp Pro Glu Ser Gly Arg Pro Arg  
3715 3720 3725

Arg Ala Gly Val Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Leu  
3730 3735 3740

Ile Leu Glu Gln Pro Pro Ser Glu Pro Ala Glu Ile Asp Gln Ser Asp  
3745 3750 3755 3760

Arg Arg Val Thr Ala His Pro Ala Val Ile Pro Trp Met Leu Ser Ala  
3765 3770 3775

Arg Ser Leu Ala Ala Leu Gln Ala Gln Ala Ala Ala Leu Gln Ala Arg  
3780 3785 3790

Leu Asp Arg Gly Pro Gly Ala Ser Pro Leu Asp Leu Gly Tyr Ser Leu  
3795 3800 3805

Ala Thr Thr Arg Ser Val Leu Asp Glu Arg Ala Val Val Trp Gly Ala  
3810 3815 3820

Asp Arg Glu Ala Leu Leu Ser Arg Leu Ala Ala Leu Ala Asp Gly Arg  
3825 3830 3835 3840

Thr Ala Pro Gly Val Ile Thr Gly Ser Ala Asn Ser Gly Gly Arg Ile  
3845 3850 3855

Gly Phe Val Phe Ser Gly Gln Gly Ser Gln Trp Leu Gly Met Gly Lys  
 3860 3865 3870

Ala Leu Cys Ala Ala Phe Pro Ala Phe Ala Asp Ala Phe Glu Glu Ala  
 3875 3880 3885

Cys Asp Ala Leu Ser Ala His Leu Gly Ala Asp Val Arg Gly Val Leu  
 3890 3895 3900

Phe Gly Ala Asp Glu Gln Met Leu Asp Arg Thr Leu Trp Ala Gln Ser  
 905 3910 3915 3920

Gly Ile Phe Ala Val Gln Val Gly Leu Leu Gly Leu Leu Arg Ser Trp  
 3925 3930 3935

Gly Val Arg Pro Ala Ala Val Leu Gly His Ser Val Gly Glu Leu Ala  
 3940 3945 3950

Ala Ala His Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Leu  
 3955 3960 3965

Val Ala Ala Arg Ala His Leu Met Gln Ala Leu Pro Thr Gly Gly Ala  
 3970 3975 3980

Met Leu Ala Val Ala Thr Ser Glu Ala Ala Val Gly Pro Leu Leu Ser  
 985 3990 3995 4000

Gly Val Cys Asp Arg Val Ser Ile Ala Ala Ile Asn Gly Pro Glu Ser  
 4005 4010 4015

Val Val Leu Ser Gly Asp Arg Asp Val Leu Val Glu Leu Ala Gly Glu  
 4020 4025 4030

Phe Asp Ala Arg Gly Leu Arg Thr Lys Trp Leu Arg Val Ser His Ala  
 4035 4040 4045

Phe His Ser His Arg Met Glu Pro Ile Leu Asp Glu Tyr Ala Glu Thr  
 4050 4055 4060

Ala Arg Cys Val Glu Phe Gly Glu Pro Val Val Pro Ile Val Ser Ala  
 065 4070 4075 4080

Ala Thr Gly Ala Leu Asp Thr Thr Gly Leu Met Cys Ala Ala Asp Tyr  
 4085 4090 4095

Trp Thr Arg Gln Val Arg Asp Pro Val Arg Phe Gly Asp Gly Val Arg  
 4100 4105 4110

Ala Leu Val Gly Gln Gly Val Asp Thr Ile Val Glu Phe Gly Pro Asp  
4115 4120 4125

Gly Ala Leu Ser Ala Leu Val Glu Gln Cys Leu Ala Gly Ser Asp Gln  
4130 4135 4140

Ala Gly Arg Val Ala Ala Ile Pro Leu Met Arg Arg Asp Arg Asp Glu  
145 4150 4155 4160

Val Glu Thr Ala Val Ala Ala Leu Ala His Val His Val Arg Gly Gly  
4165 4170 4175

Ala Val Asp Trp Ser Ala Cys Phe Ala Gly Thr Gly Ala Arg Thr Val  
4180 4185 4190

Glu Leu Pro Thr Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Ala Gly  
4195 4200 4205

Gln Ala Asp Gly Arg Gly Gly Asp Val Val Ala Asp Pro Val Asp Ala  
4210 4215 4220

Arg Phe Trp Glu Leu Val Glu Arg Ala Asp Pro Glu Pro Leu Val Asp  
225 4230 4235 4240

Glu Leu Cys Ile Asp Arg Asp Gln Pro Phe Arg Glu Val Leu Pro Val  
4245 4250 4255

Leu Ala Ser Trp Arg Glu Lys Gln Arg Gln Glu Ala Leu Ala Asp Ser  
4260 4265 4270

Trp Arg Tyr Gln Val Arg Trp Arg Ser Val Glu Val Pro Ser Ala Ala  
4275 4280 4285

Ala Leu Arg Gly Val Trp Leu Val Val Leu Pro Ala Asp Val Pro Arg  
4290 4295 4300

Asp Gln Pro Ala Val Val Ile Asp Ala Leu Ile Ala Arg Gly Ala Glu  
305 4310 4315 4320

Val Ala Val Leu Glu Leu Thr Glu Gln Asp Leu Gln Arg Ser Ala Leu  
4325 4330 4335

Val Asp Lys Val Arg Ala Val Ile Ala Asp Arg Thr Glu Val Thr Gly  
4340 4345 4350

Val Leu Ser Leu Leu Ala Met Asp Gly Met Pro Cys Ala Ala His Pro  
4355 4360 4365

His Leu Ser Arg G      al Ala Ala Thr Val Ile Leu T      ln Val Leu  
 4370                      4375                      4380

Gly Asp Ala Gly Val Ser Ala Pro Leu Trp Leu Ala Thr Thr Gly Gly  
 385                      4390                      4395                      4400

Val Glu Ala Gly Thr Glu Asp Gly Pro Ala Asp Pro Asp His Gly Leu  
                     4405                      4410                      4415

Ile Trp Gly Leu Gly Arg Val Val Gly Leu Glu His Pro Gln Trp Trp  
                     4420                      4425                      4430

Gly Gly Leu Ile Asp Leu Pro Glu Thr Leu Asp Glu Thr Ser Arg Asn  
                     4435                      4440                      4445

Gly Leu Val Ala Ala Leu Ala Gly Thr Ala Ala Glu Asp Gln Leu Ala  
                     4450                      4455                      4460

Val Arg Ser Ser Gly Leu Phe Val Arg Arg Val Val Arg Ala Ala Arg  
 465                      4470                      4475                      4480

Asn Pro Arg Ser Glu Thr Trp Arg Ser Arg Gly Thr Val Leu Ile Thr  
                     4485                      4490                      4495

Gly Gly Thr Gly Ala Leu Gly Ala Glu Val Ala Arg Trp Leu Ala Arg  
                     4500                      4505                      4510

Arg Gly Ala Glu His Leu Val Leu Ile Ser Arg Arg Gly Pro Glu Ala  
                     4515                      4520                      4525

Pro Gly Ala Ala Asp Leu Gly Ala Glu Leu Thr Glu Leu Gly Val Lys  
                     4530                      4535                      4540

Val Thr Val Leu Ala Cys Asp Val Thr Asp Arg Asp Glu Leu Ala Ala  
 4545                      4550                      4555                      4560

Val Leu Ala Ala Val Pro Thr Glu Tyr Pro Leu Ser Ala Val Val His  
                     4565                      4570                      4575

Thr Ala Gly Val Gly Thr Pro Ala Asn Leu Ala Glu Thr Thr Leu Ala  
                     4580                      4585                      4590

Gln Phe Ala Asp Val Leu Ser Ala Lys Val Val Gly Ala Ala Asn Leu  
                     4595                      4600                      4605

Asp Arg Leu Leu Gly Gly Gln Pro Leu Asp Ala Phe Val Leu Phe Ser  
                     4610                      4615                      4620



Ser Ile Ser Gly Val Trp Gly Ala Gly Gly Gln Gly Ala Tyr Ser Ala  
625 4630 4635 4640

Ala Asn Ala Tyr Leu Asp Ala Leu Ala Glu Arg Arg Arg Ala Cys Gly  
4645 4650 4655

Arg Pro Ala Thr Cys Ile Ala Trp Gly Pro Trp Ala Gly Ala Gly Met  
4660 4665 4670

Ala Val Gln Glu Gly Asn Glu Ala His Leu Arg Arg Arg Gly Leu Val  
4675 4680 4685

Pro Met Glu Pro Gln Ser Ala Leu Phe Ala Leu Gln Gln Ala Leu Ser  
4690 4695 4700

Gln Arg Glu Thr Ala Ile Thr Val Ala Asp Val Asp Trp Glu Arg Phe  
705 4710 4715 4720

Ala Ala Ser Phe Thr Ala Ala Arg Pro Arg Pro Leu Leu Glu Glu Ile  
4725 4730 4735

Val Asp Leu Arg Pro Asp Thr Glu Thr Glu Glu Lys His Gly Ala Gly  
4740 4745 4750

Glu Leu Gly Gln Gln Leu Ala Ala Leu Pro Pro Ala Glu Arg Gly His  
4755 4760 4765

Leu Leu Leu Glu Val Val Leu Ala Glu Thr Ala Ser Thr Leu Gly His  
4770 4775 4780

Asp Ser Ala Glu Ala Val Gln Pro Asp Arg Thr Phe Ala Glu Leu Gly  
785 4790 4795 4800

Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Asn Ala Val  
4805 4810 4815

Thr Gly Leu Arg Leu Pro Pro Thr Leu Val Phe Asp His Pro Thr Pro  
4820 4825 4830

Leu Ala Leu Ser Glu Gln Leu Val Pro Ala Leu Val Ala Glu Pro Asp  
4835 4840 4845

Asn Gly Ile Glu Ser Leu Leu Ala Glu Leu Asp Arg Leu Asp Thr Thr  
4850 4855 4860

Leu Ala Gln Gly Pro Ser Ile Pro Leu Glu Asp Gln Ala Lys Val Ala  
865 4870 4875 4880

Glu Arg Leu His Ala Leu Leu Ala Lys Trp Asp Gly Ala Arg Asp Gly  
 4885 4890 4895

Thr Ala Arg Ala Thr Ser Pro Gln Ser Leu Thr Ala Ala Thr Asp Asp  
 4900 4905 4910

Glu Ile Phe Asp Leu Ile Asp Arg Lys Phe Arg Arg  
 4915 4920

<210> 49  
 <211> 16767  
 <212> DNA  
 <213> Saccharopolyspora spinosa

<220>  
 <221> CDS  
 <222> (1)..(16764)  
 <223> ORF22; Polyketidsynthase

<400> 49  
 atg gcc aat gaa gaa aag ctc cgc gag tac ctc aag cgt gtc gtc gtc 48  
 Met Ala Asn Glu Glu Lys Leu Arg Glu Tyr Leu Lys Arg Val Val Val  
 1 5 10 15  
 gaa ctg gaa gag gcg cac gaa cgc ctg cac gag ttg gag cgc cag gag 96  
 Glu Leu Glu Glu Ala His Glu Arg Leu His Glu Leu Glu Arg Gln Glu  
 20 25 30  
 cac gac ccc atc gcg atc gtg tcg atg gga tgt cgt tat ccc ggt ggc 144  
 His Asp Pro Ile Ala Ile Val Ser Met Gly Cys Arg Tyr Pro Gly Gly  
 35 40 45  
 gtc tcc act ccg gag gag ctg tgg cga ctg gtc gtc gac gga gga gac 192  
 Val Ser Thr Pro Glu Glu Leu Trp Arg Leu Val Val Asp Gly Gly Asp  
 50 55 60  
 gcg atc gcg aac ttc ccc gaa gac cgt ggc tgg aat ctg gac gag ctg 240  
 Ala Ile Ala Asn Phe Pro Glu Asp Arg Gly Trp Asn Leu Asp Glu Leu  
 65 70 75 80  
 ttc gat cct gat ccg ggc cga gcc ggg acc tcc tac gtc cgc gag ggt 288  
 Phe Asp Pro Asp Pro Gly Arg Ala Gly Thr Ser Tyr Val Arg Glu Gly  
 85 90 95  
 ggt ttc ctg cgc ggg gtc gcg gac ttc gat gcc ggg ctc ttc ggg atc 336  
 Gly Phe Leu Arg Gly Val Ala Asp Phe Asp Ala Gly Leu Phe Gly Ile

100	105	110	
agt ccg cgc gag gca cag gcg atg gac ccg caa cag cgg ttg ctg ctg			384
Ser Pro Arg Glu Ala Gln Ala Met Asp Pro Gln Gln Arg Leu Leu Leu			
115	120	125	
gag atc tcg tgg gag gtg ttc gag cgc gcc ggc att gac ccg ttt tct			432
Glu Ile Ser Trp Glu Val Phe Glu Arg Ala Gly Ile Asp Pro Phe Ser			
130	135	140	
ttg cgg ggt acc aag acc ggt gtg ttc gcg ggc ctg atc tac cac gac			480
Leu Arg Gly Thr Lys Thr Gly Val Phe Ala Gly Leu Ile Tyr His Asp			
145	150	155	160
tac gcg tcg cgg ttt cgc aag acc ccc gcg gag ttc gag ggt tac ttc			528
Tyr Ala Ser Arg Phe Arg Lys Thr Pro Ala Glu Phe Glu Gly Tyr Phe			
165	170	175	
gcc acc ggc aac gcg ggc agc gtc gca tcc ggc cgg gtg gct tac acc			576
Ala Thr Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Thr			
180	185	190	
ttc ggg tta gag ggc ccg gcg gtc acc gtg gac acc gcc tgc tcg tcg			624
Phe Gly Leu Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser			
195	200	205	
tcc ctg gtg gcg ctg cac ctg gcc tgc cag tcc ctg cgg ctg ggc gaa			672
Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Leu Gly Glu			
210	215	220	
tgc gac ctg gcc ctg gcc ggt ggc att tcg gtg atg gcc acg ccg gga			720
Cys Asp Leu Ala Leu Ala Gly Gly Ile Ser Val Met Ala Thr Pro Gly			
225	230	235	240
gcc ttc gtc gag ttc agc cgg caa cgc gca ctc gcc tcg gat ggc cgg			768
Ala Phe Val Glu Phe Ser Arg Gln Arg Ala Leu Ala Ser Asp Gly Arg			
245	250	255	
tgc aag ccc ttc gcg gat gcc gcc gac ggc acc ggc tgg ggc gag ggc			816
Cys Lys Pro Phe Ala Asp Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly			
260	265	270	
gcc gga atg ctg ctg ctg gaa cgg ctg tcg gac gca cga cga aac ggc			864
Ala Gly Met Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly			
275	280	285	
cac ccg gtg ctg gcg gcg gtg gtc ggt tcc gcg atc aac cag gac ggg			912
His Pro Val Leu Ala Ala Val Val Gly Ser Ala Ile Asn Gln Asp Gly			

290	295	300	
acg tcc aac ggc ctg acc gcg ccc agc ggt ccc gca cag cag cga gtg			960
Thr Ser Asn Gly Leu Thr Ala Pro Ser Gly Pro Ala Gln Gln Arg Val			
305	310	315	320
atc cgc caa gcc ctg gcg aac gcc ggg ttg tcg ccc gcc gag gtc gat			1008
Ile Arg Gln Ala Leu Ala Asn Ala Gly Leu Ser Pro Ala Glu Val Asp			
325	330	335	
gtg gtc gag gcg cac ggc acg ggc acg gcc ttg ggc gac ccg atc gag			1056
Val Val Glu Ala His Gly Thr Gly Thr Ala Leu Gly Asp Pro Ile Glu			
340	345	350	
gcg cag gcc ctg atc gcc acc tac ggg gcg aac cgg tcg gcg gat cat			1104
Ala Gln Ala Leu Ile Ala Thr Tyr Gly Ala Asn Arg Ser Ala Asp His			
355	360	365	
ccg ctg ctg ctg ggt tcc ctc aag tcg aac atc ggc cac acc cag gct			1152
Pro Leu Leu Leu Gly Ser Leu Lys Ser Asn Ile Gly His Thr Gln Ala			
370	375	380	
gcc gcc ggt gtg gcc ggg gtg atc aag tcg gtc ctg gcc atc agg cac			1200
Ala Ala Gly Val Ala Gly Val Ile Lys Ser Val Leu Ala Ile Arg His			
385	390	395	400
cgg gag atg ccc cgc agc ctg cac atc gac cag cca tcg cag cac gtg			1248
Arg Glu Met Pro Arg Ser Leu His Ile Asp Gln Pro Ser Gln His Val			
405	410	415	
gac tgg tcg gcg ggc gcg gtg cgg ctg ctc acg gac agc gtt gac tgg			1296
Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Asp Ser Val Asp Trp			
420	425	430	
ccg gat ctc ggc agg ccg cgc cga gca ggg gtg tcc tcg ttc ggc atg			1344
Pro Asp Leu Gly Arg Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Met			
435	440	445	
agc ggt acc aac gca cac ctg atc gtc gag gaa gta tcc gac gag ccg			1392
Ser Gly Thr Asn Ala His Leu Ile Val Glu Glu Val Ser Asp Glu Pro			
450	455	460	
gtc tcg ggc agt acc gag ccg acc ggg gca ttt ccc tgg ccg ctg tcc			1440
Val Ser Gly Ser Thr Glu Pro Thr Gly Ala Phe Pro Trp Pro Leu Ser			
465	470	475	480
ggc aag acg gag acg gca ttg cgc gag cag gct gcc gag ttg ctc tcc			1488
Gly Lys Thr Glu Thr Ala Leu Arg Glu Gln Ala Ala Glu Leu Leu Ser			

485	490	495	
gta gtg acc gag cac ccg gag ccg gga ctg ggg gac gtc ggg tac tcg			1536
Val Val Thr Glu His Pro Glu Pro Gly Leu Gly Asp Val Gly Tyr Ser			
500	505	510	
ctg gcc acc ggt cgc gct gcg atg gag cac ccg gct gtc gtg gtt gcc			1584
Leu Ala Thr Gly Arg Ala Ala Met Glu His Arg Ala Val Val Val Ala			
515	520	525	
gac gat ccg gac tct ttc gtc gcc gga ctg acg gcg ttg gct gcg ggc			1632
Asp Asp Arg Asp Ser Phe Val Ala Gly Leu Thr Ala Leu Ala Ala Gly			
530	535	540	
gtt ccg gca gcc aac gtg gtg cag ggc gcg gcc gac tgc aag gga aag			1680
Val Pro Ala Ala Asn Val Val Gln Gly Ala Ala Asp Cys Lys Gly Lys			
545	550	555	560
gtc gcg ttc gtg ttc ccc ggc cag ggc tcg cat tgg cag ggg atg gcg			1728
Val Ala Phe Val Phe Pro Gly Gln Gly Ser His Trp Gln Gly Met Ala			
565	570	575	
agg gaa ctg tcc gaa tcc tcg ccg gtg ttc ccg ccg aag ctg gcg gaa			1776
Arg Glu Leu Ser Glu Ser Ser Pro Val Phe Arg Arg Lys Leu Ala Glu			
580	585	590	
tgc gcg gcg gct acg gcc cct tac gtg gac tgg tcg ctg ctc ggc gtc			1824
Cys Ala Ala Ala Thr Ala Pro Tyr Val Asp Trp Ser Leu Leu Gly Val			
595	600	605	
ctt cgc ggt gat ccc gat gca ccc gcg ctg gat cgc gac gac gtg att			1872
Leu Arg Gly Asp Pro Asp Ala Pro Ala Leu Asp Arg Asp Asp Val Ile			
610	615	620	
cag ctc gcg ctg ttc gcc atg atg gtg tcg ctg gcc gaa ctg tgg cgt			1920
Gln Leu Ala Leu Phe Ala Met Met Val Ser Leu Ala Glu Leu Trp Arg			
625	630	635	640
tcg tgc gga gtg gag ccc gcc gcg gtg gtc ggt cat tcc cag ggc gag			1968
Ser Cys Gly Val Glu Pro Ala Ala Val Val Gly His Ser Gln Gly Glu			
645	650	655	
atc gcc gcc gcc cat gtg gca ggc gct ttg tcc ttg act gat gcg gtg			2016
Ile Ala Ala Ala His Val Ala Gly Ala Leu Ser Leu Thr Asp Ala Val			
660	665	670	
cgc atc atc gct gcc cgc tgc gat gcg gtg tcg gcg ctg acc ggg aag			2064
Arg Ile Ile Ala Ala Arg Cys Asp Ala Val Ser Ala Leu Thr Gly Lys			

gga ggc atg ctc gcg att gcc ttg ccg gaa agc gcg gtg gtg aag cga 2112  
 Gly Gly Met Leu Ala Ile Ala Leu Pro Glu Ser Ala Val Val Lys Arg  
 690 695 700

atc gca ggc ctg ccg gag ctg acc gtt gcg gcg gtc aac gga ccc ggc 2160  
 Ile Ala Gly Leu Pro Glu Leu Thr Val Ala Ala Val Asn Gly Pro Gly  
 705 710 715 720

tcc act gtc gtt tcc ggc gaa ccg tcg gct ctg gag cgt ctg cag acc 2208  
 Ser Thr Val Val Ser Gly Glu Pro Ser Ala Leu Glu Arg Leu Gln Thr  
 725 730 735

gaa ctg acc gcg gaa aac gtg cag acc ccg ccg gtg gga att gat tac 2256  
 Glu Leu Thr Ala Glu Asn Val Gln Thr Arg Arg Val Gly Ile Asp Tyr  
 740 745 750

gcc tcg cat tcg ccg cag atc gcg cag gtc cag ggc ccg ctt ctg gac 2304  
 Ala Ser His Ser Pro Gln Ile Ala Gln Val Gln Gly Arg Leu Leu Asp  
 755 760 765

cgg ctg ggc gaa gtc ggg tcc gaa cct gct gag atc gct ttc tac tcg 2352  
 Arg Leu Gly Glu Val Gly Ser Glu Pro Ala Glu Ile Ala Phe Tyr Ser  
 770 775 780

acg gtc acc ggc gag ccg acc gac acc ggc cga ctc gac gcc gac tac 2400  
 Thr Val Thr Gly Glu Arg Thr Asp Thr Gly Arg Leu Asp Ala Asp Tyr  
 785 790 795 800

tgg tac cag aac ctt ccg cag ccc gtc cgc ttc cag cag acc gtc gcc 2448  
 Trp Tyr Gln Asn Leu Arg Gln Pro Val Arg Phe Gln Gln Thr Val Ala  
 805 810 815

cgg atg gca gat cag ggc tat ccg ttc ttc gtc gag gtg agc ccg cac 2496  
 Arg Met Ala Asp Gln Gly Tyr Arg Phe Phe Val Glu Val Ser Pro His  
 820 825 830

ccg ctg ctc acc gcc gga atc cag gaa acg ctg gaa gcc gcg gac gcg 2544  
 Pro Leu Leu Thr Ala Gly Ile Gln Glu Thr Leu Glu Ala Ala Asp Ala  
 835 840 845

ggc ggg gtg gtg gtc ggt tcg ctg ccg cgt ggc gag ggc ggc tcc ccg 2592  
 Gly Gly Val Val Val Gly Ser Leu Arg Arg Gly Glu Gly Gly Ser Arg  
 850 855 860

cgc tgg ctg act tcg ctg gcc gag tgc cag gtg cgc gga ctg ccg gtg 2640  
 Arg Trp Leu Thr Ser Leu Ala Glu Cys Gln Val Arg Gly Leu Pro Val

865

870

875

880

aat tgg gaa cag gta ttc ctc aac acc gga gcc cga cgc gtg ccg ctg 2688  
 Asn Trp Glu Gln Val Phe Leu Asn Thr Gly Ala Arg Arg Val Pro Leu  
 885 890 895

ccg acc tac ccg ttc cag cgg cag cgg tac tgg ttg gag tcc gcc gag 2736  
 Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Glu Ser Ala Glu  
 900 905 910

tac gac gcg ggc gat ctc ggt tcg gtg ggc ttg ctc tcc gcc gag cat 2784  
 Tyr Asp Ala Gly Asp Leu Gly Ser Val Gly Leu Leu Ser Ala Glu His  
 915 920 925

ccc ctg ctc ggg gct gcg gtg acg ctg gcc gat gcg ggc ggg ttc ctg 2832  
 Pro Leu Leu Gly Ala Ala Val Thr Leu Ala Asp Ala Gly Gly Phe Leu  
 930 935 940

ctg acc ggc aag ctg tcg gtc aag acc cag ccc tgg ttg gcc gac cac 2880  
 Leu Thr Gly Lys Leu Ser Val Lys Thr Gln Pro Trp Leu Ala Asp His  
 945 950 955 960

gtg gtc ggc ggg gcg atc ctg ctg ccc ggc acc gcg ttc gtg gaa atg 2928  
 Val Val Gly Gly Ala Ile Leu Leu Pro Gly Thr Ala Phe Val Glu Met  
 965 970 975

ctg ata cgc gcc gcg gac cag gtc ggg tgc gat ctg atc gag gag ttg 2976  
 Leu Ile Arg Ala Ala Asp Gln Val Gly Cys Asp Leu Ile Glu Glu Leu  
 980 985 990

tcc ctg acg act ccg ctg gtt ttg ccc gcg acc ggt gcg gtg cag gtg 3024  
 Ser Leu Thr Thr Pro Leu Val Leu Pro Ala Thr Gly Ala Val Gln Val  
 995 1000 1005

cag atc gcg gtt ggc ggt ccg gac gag gcc ggg cgc cgc tcg gtc cgc 3072  
 Gln Ile Ala Val Gly Gly Pro Asp Glu Ala Gly Arg Arg Ser Val Arg  
 1010 1015 1020

gtg cat tcc tgt cga gac gac gcc gtg ccg cag gac tcg tgg acc tgc 3120  
 Val His Ser Cys Arg Asp Asp Ala Val Pro Gln Asp Ser Trp Thr Cys  
 1025 1030 1035 1040

cac gcg acc ggc acg ttg acc tcc agc gat cac cag gac gcc ggc cag 3168  
 His Ala Thr Gly Thr Leu Thr Ser Ser Asp His Gln Asp Ala Gly Gln  
 1045 1050 1055

ggc ccc gat ggg att tgg ccg ccc aac gat gct gtc gcg gtt ccg ctg 3216  
 Gly Pro Asp Gly Ile Trp Pro Pro Asn Asp Ala Val Ala Val Pro Leu

1060

1065

1070

gac agc ttc tac gcc cgc gca gct gag cgg ggc ttc gat ttc ggc ccg 3264  
 Asp Ser Phe Tyr Ala Arg Ala Ala Glu Arg Gly Phe Asp Phe Gly Pro  
 1075 1080 1085

gcg ttc cag ggg ttg cag gcg gct tgg aag cgc gga gac gag atc ttc 3312  
 Ala Phe Gln Gly Leu Gln Ala Ala Trp Lys Arg Gly Asp Glu Ile Phe  
 1090 1095 1100

gcc gag gtc ggc ctg ccc acc gca cac cgc gaa gac gcc ggc agg ttc 3360  
 Ala Glu Val Gly Leu Pro Thr Ala His Arg Glu Asp Ala Gly Arg Phe  
 1105 1110 1115 1120

gga atc cac cct gct ctg ctg gat gcg gca ctg cag gcg ctg ggc gca 3408  
 Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu Gln Ala Leu Gly Ala  
 1125 1130 1135

gcc gaa gag gat ccg gac gag gga tgg ctc ccg ttc gcg tgg caa ggt 3456  
 Ala Glu Glu Asp Pro Asp Glu Gly Trp Leu Pro Phe Ala Trp Gln Gly  
 1140 1145 1150

gtg tcc ctc aaa gcg acg ggc gca ctt tcc ctt cgg gtg cac ctc gtt 3504  
 Val Ser Leu Lys Ala Thr Gly Ala Leu Ser Leu Arg Val His Leu Val  
 1155 1160 1165

ccg gcg ggc gcg aat gcg gtg tcg gtg ttc acg acc gac acg act ggc 3552  
 Pro Ala Gly Ala Asn Ala Val Ser Val Phe Thr Thr Asp Thr Thr Gly  
 1170 1175 1180

caa gcc gtg ctc tcc atc gat tcg ctg gtg ctg cgc cag att tcg gac 3600  
 Gln Ala Val Leu Ser Ile Asp Ser Leu Val Leu Arg Gln Ile Ser Asp  
 1185 1190 1195 1200

aag cag ttg gca gcg gcc cgt gcg atg gaa cac gag tcc ctg ttc ccg 3648  
 Lys Gln Leu Ala Ala Ala Arg Ala Met Glu His Glu Ser Leu Phe Arg  
 1205 1210 1215

gtc gac tgg aag cga atc tcg ccc ggc gct gcc aag ccg gtc tcc tgg 3696  
 Val Asp Trp Lys Arg Ile Ser Pro Gly Ala Ala Lys Pro Val Ser Trp  
 1220 1225 1230

gca gtg atc ggc aat gac gaa ctc gcc cga gcc tgc ggc tcg gca ctt 3744  
 Ala Val Ile Gly Asn Asp Glu Leu Ala Arg Ala Cys Gly Ser Ala Leu  
 1235 1240 1245

ggc acg gaa ctc cac ccc gac ctg acc ggg ttg gct gac ccg ccc ccg 3792  
 Gly Thr Glu Leu His Pro Asp Leu Thr Gly Leu Ala Asp Pro Pro Pro



1250	1255	1260	
gac gtc gtg gtg gtg cca tgc ggt gcg tct cgc cag gac ttg gac gtt			3840
Asp Val Val Val Val Pro Cys Gly Ala Ser Arg Gln Asp Leu Asp Val			
1265	1270	1275	1280
gct tcc gag gca cgt gcc gcg aca caa cgc atg ctt gac ctg atc cag			3888
Ala Ser Glu Ala Arg Ala Ala Thr Gln Arg Met Leu Asp Leu Ile Gln			
1285	1290	1295	
gat tgg ttg gcg gcg gcg cga ttc gcc gga tct cgc ctg gtg gtt gtg			3936
Asp Trp Leu Ala Ala Ala Arg Phe Ala Gly Ser Arg Leu Val Val Val			
1300	1305	1310	
acg tgt ggt gcg gcg tcg aca ggt ccc gcc gag ggt gtt tcc gac ctg			3984
Thr Cys Gly Ala Ala Ser Thr Gly Pro Ala Glu Gly Val Ser Asp Leu			
1315	1320	1325	
gtg cat gct gcg tcg tgg ggt ttg ttg cgt tcg gcg cag tcg gag aac			4032
Val His Ala Ala Ser Trp Gly Leu Leu Arg Ser Ala Gln Ser Glu Asn			
1330	1335	1340	
ccg gac cga ttc gtg ttg gtc gat gtg gac gga acc gcc gaa tca tgg			4080
Pro Asp Arg Phe Val Leu Val Asp Val Asp Gly Thr Ala Glu Ser Trp			
1345	1350	1355	1360
cgt gcg ctc gcg gcg gcc gtg cgt tcc gga gaa ccg cag ctg gcg ttg			4128
Arg Ala Leu Ala Ala Ala Val Arg Ser Gly Glu Pro Gln Leu Ala Leu			
1365	1370	1375	
cgc gcc ggt gaa gtc cgg gtg cct cgc ctg gcg cga tgt gtt gcc gcc			4176
Arg Ala Gly Glu Val Arg Val Pro Arg Leu Ala Arg Cys Val Ala Ala			
1380	1385	1390	
gag gac agc cgg atc cca gtg ccc ggt gcg gat ggg acg gtg ttg att			4224
Glu Asp Ser Arg Ile Pro Val Pro Gly Ala Asp Gly Thr Val Leu Ile			
1395	1400	1405	
tcc ggc ggt acg ggc ctg ctg ggc ggg ttg gtt gcc cgg cat ttg gtg			4272
Ser Gly Gly Thr Gly Leu Leu Gly Gly Leu Val Ala Arg His Leu Val			
1410	1415	1420	
gcg gag cgc ggt gtc cgc cgc ctg gtg ctc gcg ggg cga cgc ggc tgg			4320
Ala Glu Arg Gly Val Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp			
1425	1430	1435	1440
agc gcc ccc ggg gtc acc gac ctg gtg gat gag ttg gtg ggc ctg gga			4368
Ser Ala Pro Gly Val Thr Asp Leu Val Asp Glu Leu Val Gly Leu Gly			

1445

1450

1455

gct gcg gtc gag gtg gcg agc tgc gat gtc ggg gat cgg gcc cag ttg 4416  
 Ala Ala Val Glu Val Ala Ser Cys Asp Val Gly Asp Arg Ala Gln Leu  
 1460 1465 1470

gac cgg ctg ctg acg acg atc tcg gca gag ttc ccg ctg cgc gga gtg 4464  
 Asp Arg Leu Leu Thr Thr Ile Ser Ala Glu Phe Pro Leu Arg Gly Val  
 1475 1480 1485

gtg cat gcg gcc ggg gca ctt gcc gac ggg gtc gtc gag tcg ctg aca 4512  
 Val His Ala Ala Gly Ala Leu Ala Asp Gly Val Val Glu Ser Leu Thr  
 1490 1495 1500

cca gag cac gtg gca aag gtg ttc ggc ccg aag gcc gcc ggt gcg tgg 4560  
 Pro Glu His Val Ala Lys Val Phe Gly Pro Lys Ala Ala Gly Ala Trp  
 1505 1510 1515 1520

cac ctg cac gag ttg act ctt gat ctg gat ctc tcg ttc ttc gtg ctc 4608  
 His Leu His Glu Leu Thr Leu Asp Leu Asp Leu Ser Phe Phe Val Leu  
 1525 1530 1535

ttc tcc tcg ttc tcc ggc gtg gcg ggg gct gcg ggt cag gga aac tac 4656  
 Phe Ser Ser Phe Ser Gly Val Ala Gly Ala Ala Gly Gln Gly Asn Tyr  
 1540 1545 1550

gcg gcg gcg aac gcg ttc ctg gac ggc ctg gct cag cac cgg cgg acg 4704  
 Ala Ala Ala Asn Ala Phe Leu Asp Gly Leu Ala Gln His Arg Arg Thr  
 1555 1560 1565

gcg ggg ctg cct gcg gtg tcg ctg gct tgg ggc ttg tgg gag cag ccc 4752  
 Ala Gly Leu Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro  
 1570 1575 1580

agc ggg atg acc gga gcg ctc gat gcg gcg ggc cgt agc cgc att gcg 4800  
 Ser Gly Met Thr Gly Ala Leu Asp Ala Ala Gly Arg Ser Arg Ile Ala  
 1585 1590 1595 1600

cgc acc aat ccg ccg atg tcc gcg ccg gac ggg ttg cgg ctg ttc gag 4848  
 Arg Thr Asn Pro Pro Met Ser Ala Pro Asp Gly Leu Arg Leu Phe Glu  
 1605 1610 1615

atg gcg ttt cgc gtt ccg ggc gaa tcg ctt ctg gtt ccg gtc cac gtc 4896  
 Met Ala Phe Arg Val Pro Gly Glu Ser Leu Leu Val Pro Val His Val  
 1620 1625 1630

gac ctg aac gcc ctg cgc gct gat gcg gcc gac ggc ggt gtg cct gcg 4944  
 Asp Leu Asn Ala Leu Arg Ala Asp Ala Ala Asp Gly Gly Val Pro Ala

1635	1640	1645	
ttg ttg cgc gac ctg gtg cca gcg ccc gtg cgg cgg agc gcg gtc aac			4992
Leu Leu Arg Asp Leu Val Pro Ala Pro Val Arg Arg Ser Ala Val Asn			
1650	1655	1660	
gag tcg gcg gac gtc aac ggt ctg gtt ggt cgg ctg cgg agg ctg ccg			5040
Glu Ser Ala Asp Val Asn Gly Leu Val Gly Arg Leu Arg Arg Leu Pro			
1665	1670	1675	1680
gac ctg gat cag gaa acc cag ctg ttg ggt ttg gtg cgc gag cat gtt			5038
Asp Leu Asp Gln Glu Thr Gln Leu Leu Gly Leu Val Arg Glu His Val			
1685	1690	1695	
tcg gcg gtg ctg ggg cat tcg ggt gcg gtc gag gtc ggg gcc gat cgt			5136
Ser Ala Val Leu Gly His Ser Gly Ala Val Glu Val Gly Ala Asp Arg			
1700	1705	1710	
gct ttc cgg gat ttg ggt ttt gat tcg ttg tcc ggt gtg gag ttt cgg			5184
Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Phe Arg			
1715	1720	1725	
aac cgg ctt ggc ggg gtg ctg ggc gtt cgg ttg ccg gct act gcg gtg			5232
Asn Arg Leu Gly Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val			
1730	1735	1740	
ttc gac tat ccg aca ccg cgg gcg ttg gtt cgg ttc ttg ctc gac aaa			5280
Phe Asp Tyr Pro Thr Pro Arg Ala Leu Val Arg Phe Leu Leu Asp Lys			
1745	1750	1755	1760
ctg att ggt ggc gtg gag gct ccg act ccc gca ccg gcg gct gtg gcg			5328
Leu Ile Gly Gly Val Glu Ala Pro Thr Pro Ala Pro Ala Ala Val Ala			
1765	1770	1775	
gcg gtg act gct gac gat ccc gtt gtg atc gtg ggg atg ggc tgt cgt			5376
Ala Val Thr Ala Asp Asp Pro Val Val Ile Val Gly Met Gly Cys Arg			
1780	1785	1790	
tat ccg ggt ggg gtg tcc tcg ccg gag gag ctt tgg cgt ttg gtg gcc			5424
Tyr Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala			
1795	1800	1805	
ggg ggc ttg gat gcg gtg gcg gag ttc ccg gac gat cgt ggc tgg gat			5472
Gly Gly Leu Asp Ala Val Ala Glu Phe Pro Asp Asp Arg Gly Trp Asp			
1810	1815	1820	
cag gcg ggg ttg ttc gat ccg gat ccc gat cgt ctt ggg acc tcg tat			5520
Gln Ala Gly Leu Phe Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr			

1825	1830	1835	1840	
gtg tgt gag ggt ggc ttc ctg cga gat gcg gca gag ttc gat gcc ggt				5568
Val Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Gly				
1845	1850	1855		
ttc ttc ggg att tcc ccg cgt gag gcg ttg gcg atg gat ccg cag cag				5616
Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln				
1860	1865	1870		
cgg ttg ctg ctg gaa gtc gct tgg gaa acc gtg gag cgg gcg ggg att				5664
Arg Leu Leu Leu Glu Val Ala Trp Glu Thr Val Glu Arg Ala Gly Ile				
1875	1880	1885		
gat ccg ctt tcg ttg cgg ggg agc cgg acc ggc gtg ttc gcg ggg ctg				5712
Asp Pro Leu Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu				
1890	1895	1900		
atg cac cac gac tac ggc gcg cgg ttc atc acg agg gcg ccg gag ggt				5760
Met His His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly				
1905	1910	1915	1920	
ttc gag ggt tat cta ggt aat ggc agc gcg gga ggc gtg ttt tcg ggt				5808
Phe Glu Gly Tyr Leu Gly Asn Gly Ser Ala Gly Gly Val Phe Ser Gly				
1925	1930	1935		
cgg gtt gcg tat tcg ttt ggt ttc gag ggt cct gcg gtg acg gtg gat				5856
Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp				
1940	1945	1950		
acg gcg tgt tcg tcg tcg ttg gtg gcg ctg cac ctg gcg ggt caa gca				5904
Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala				
1955	1960	1965		
ctg cgg tct ggt gag tgt gat ctg gct ctt gcg ggt ggt gtg acg gtg				5952
Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val				
1970	1975	1980		
atg gcc acg ccg ggg atg ttc gtg gag ttt tcg cgt caa cgg ggc ttg				6000
Met Ala Thr Pro Gly Met Phe Val Glu Phe Ser Arg Gln Arg Gly Leu				
1985	1990	1995	2000	
gcg gcg gat ggg cgg tgc aag tcg ttt gcg gcg gct gcg gat ggc acc				6048
Ala Ala Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Asp Gly Thr				
2005	2010	2015		
ggt tgg gga gaa ggc gcg ggc ttg gtg ttg ttg gag cgg ctg tcg gat				6096
Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp				

2020	2025	2030	
gcc cgg cgc aac ggg cac gcg gtt ctg gcg gtc gtg cgg ggt agc gcg Ala Arg Arg Asn Gly His Ala Val Leu Ala Val Val Arg Gly Ser Ala 2035 2040 2045			6144
gtg aat cag gat ggt gcg tcg aat ggt ttg acg gcg ccg aat ggg ccc Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro 2050 2055 2060			6192
tcg cag cag cgg gtg atc acg cag gcg ttg gcg agt gct ggt ttg tcg Ser Gln Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser 2065 2070 2075 2080			6240
gtg tct gat gtg gac gcc gtg gag gcg cat ggg act gga acc agg ctt Val Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu 2085 2090 2095			6288
ggt gat ccg att gag gcg cag gct ctg att gcc act tac ggg cag ggg Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly 2100 2105 2110			6336
cgg gat agc gat cgg ccg ttg tgg ttg ggg tcg gtg aag tcg aat att Arg Asp Ser Asp Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile 2115 2120 2125			6384
ggt cat acg cag gcg gcg gcg ggt gtc gct ggt gtg atc aag atg gtg Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val 2130 2135 2140			6432
atg gcg atg cgg cac ggg cag ctg ccc gcg acg ttg cat gtg gat gaa Met Ala Met Arg His Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu 2145 2150 2155 2160			6480
cct acg tcg gaa gtg gat tgg tcg gcg ggg gat gtc cag ctc ctc acg Pro Thr Ser Glu Val Asp Trp Ser Ala Gly Asp Val Gln Leu Leu Thr 2165 2170 2175			6528
gag aac acc ccc tgg ccc ggc aac agc cat cct cgg cgg gtg ggc gtg Glu Asn Thr Pro Trp Pro Gly Asn Ser His Pro Arg Arg Val Gly Val 2180 2185 2190			6576
tcg tcg ttc ggg atc agc ggc acc aac gca cac gtc atc ctc gaa caa Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln 2195 2200 2205			6624
gcc tcg aaa aca cca gac gag act gcg gac aag agc ggt ccc gat tcg Ala Ser Lys Thr Pro Asp Glu Thr Ala Asp Lys Ser Gly Pro Asp Ser			6672

2210	2215	2220	
gaa tcg acc gtg gac ctt cca gcg gtc ccg ttg atc gtg tcg ggg aga			6720
Glu Ser Thr Val Asp Leu Pro Ala Val Pro Leu Ile Val Ser Gly Arg			
2225	2230	2235	2240
aca ccg gca gcg ctc agc gct cag gcg agc gca ttg ttg tcc tat ttg			6768
Thr Pro Ala Ala Leu Ser Ala Gln Ala Ser Ala Leu Leu Ser Tyr Leu			
2245	2250	2255	
ggc gag cgt ggc gat att tcc acg ctg gat gcg gcg ttt tcg ttg gct			6816
Gly Glu Arg Gly Asp Ile Ser Thr Leu Asp Ala Ala Phe Ser Leu Ala			
2260	2265	2270	
tcc tcc ccg gcc gcg ttg gag gag ccg gcg gtg gtg ctg gga gcg gac			6864
Ser Ser Arg Ala Ala Leu Glu Glu Arg Ala Val Val Leu Gly Ala Asp			
2275	2280	2285	
cgc gaa acg ttg ttg tcc ggg ttg gaa gcg ctg gct tcc ggt cgc gag			6912
Arg Glu Thr Leu Leu Ser Gly Leu Glu Ala Leu Ala Ser Gly Arg Glu			
2290	2295	2300	
gct tct ggg gtg gtg tcg gga tcc ccg gtc tct ggc ggg gtt ggg ttc			6960
Ala Ser Gly Val Val Ser Gly Ser Pro Val Ser Gly Gly Val Gly Phe			
2305	2310	2315	2320
gtg ttc gcc ggt cag ggc gga cag tgg ttg ggg atg ggc cgg ggg ctc			7008
Val Phe Ala Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu			
2325	2330	2335	
tac tcg gtt ttt ccg gtg ttc gct gac gcg ttt gac gaa gca tgt gcc			7056
Tyr Ser Val Phe Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala			
2340	2345	2350	
gga ctg gac gcg cat ctg ggg cag gac gtg ggg gtc cgg gat gtg gtg			7104
Gly Leu Asp Ala His Leu Gly Gln Asp Val Gly Val Arg Asp Val Val			
2355	2360	2365	
ttt ggt tcc gac ggg tcc ttg ttg gat ccg acg ctg tgg gcc cag tcg			7152
Phe Gly Ser Asp Gly Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser			
2370	2375	2380	
ggc ttg ttc gcg ttg cag gtt ggt ttg ctg agc ctg ctg ggt tcg tgg			7200
Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Ser Leu Leu Gly Ser Trp			
2385	2390	2395	2400
ggc gtc ccg ccg ggt gtg gtg ctg ggc cat tcg gtc ggc gag ttc gcg			7248
Gly Val Arg Pro Gly Val Val Leu Gly His Ser Val Gly Glu Phe Ala			

2405	2410	2415	
gcg gcg gtt gcg gcg gga gtg ttg tcg ttg ccg gat gcg gct cgg atg			7296
Ala Ala Val Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Met			
2420	2425	2430	
gtg gcg ggt cgt gcc cgg ttg atg cag gcg ttg cct tct ggc ggt gcc			7344
Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala			
2435	2440	2445	
atg ttg gcg gtg gct gct ggt gag gag cag ctg cgg ccg ttg ttg gcc			7392
Met Leu Ala Val Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala			
2450	2455	2460	
gat cgg gtt gat ggt gcg ggt atc gcc gcg gtc aac gct cct gag tcg			7440
Asp Arg Val Asp Gly Ala Gly Ile Ala Ala Val Asn Ala Pro Glu Ser			
2465	2470	2475	2480
gtg gtg ctc tcc ggc gat cgg gag gtg ctt gac gac atc gcc ggc gcg			7488
Val Val Leu Ser Gly Asp Arg Glu Val Leu Asp Asp Ile Ala Gly Ala			
2485	2490	2495	
ctg gat ggg caa ggg att cgg tgg cgg cgg ttg cgg gtt tcg cat gcg			7536
Leu Asp Gly Gln Gly Ile Arg Trp Arg Arg Leu Arg Val Ser His Ala			
2500	2505	2510	
ttt cat tcg tat cgg atg gac ccg atg ttg cag gag ttc gcc gaa atc			7584
Phe His Ser Tyr Arg Met Asp Pro Met Leu Gln Glu Phe Ala Glu Ile			
2515	2520	2525	
gca cgc agc gtg gac tac cgg cgt ggc gac cta ccg gtc gtg tcg acg			7632
Ala Arg Ser Val Asp Tyr Arg Arg Gly Asp Leu Pro Val Val Ser Thr			
2530	2535	2540	
ttg acg ggt gag ctc gac acc gca ggt gtg atg gct acg ccg gag tat			7680
Leu Thr Gly Glu Leu Asp Thr Ala Gly Val Met Ala Thr Pro Glu Tyr			
2545	2550	2555	2560
tgg gtg cgt cag gtt cga gag ccc gtc cgc ttc gcc gac ggc gtc cgg			7728
Trp Val Arg Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Arg			
2565	2570	2575	
gtg ctc gcg cag caa ggg gtc gcc acg atc ttc gaa ctc ggc cct gat			7776
Val Leu Ala Gln Gln Gly Val Ala Thr Ile Phe Glu Leu Gly Pro Asp			
2580	2585	2590	
gcg acg ctg tcg gcc ctg att ccc gat tgt cat tcg tgg gct gat cag			7824
Ala Thr Leu Ser Ala Leu Ile Pro Asp Cys His Ser Trp Ala Asp Gln			

2595

2600

2605

gcc atg ccg att ccg atg ctg cgt aaa gac cgt acg gaa acc gaa act 7872  
 Ala Met Pro Ile Pro Met Leu Arg Lys Asp Arg Thr Glu Thr Glu Thr  
 2610 2615 2620

gtg gtc gcc gcg gtg gcg cgg gcg cac acg cgt ggt gtt ccg gtc gaa 7920  
 Val Val Ala Ala Val Ala Arg Ala His Thr Arg Gly Val Pro Val Glu  
 2625 2630 2635 2640

tgg tcg gcg tat ttc gcc ggc acc ggg gca cgg cgg gtc gag ttg ccg 7968  
 Trp Ser Ala Tyr Phe Ala Gly Thr Gly Ala Arg Arg Val Glu Leu Pro  
 2645 2650 2655

acg tat gcc ttc cag cgg cag cgg tac tgg ctg gaa aca tcg gat tac 8016  
 Thr Tyr Ala Phe Gln Arg Gln Arg Tyr Trp Leu Glu Thr Ser Asp Tyr  
 2660 2665 2670

ggc gat gtg acg ggt atc ggc ctg gct gcg gcg gag cat ccg ttg ctg 8064  
 Gly Asp Val Thr Gly Ile Gly Leu Ala Ala Ala Glu His Pro Leu Leu  
 2675 2680 2685

ggg gcc gtg gtt gcg ctg gcc gat ggt gat ggg atg gtg ctg acc ggc 8112  
 Gly Ala Val Val Ala Leu Ala Asp Gly Asp Gly Met Val Leu Thr Gly  
 2690 2695 2700

cgg ttg tcg gtg ggg acg cat ccg tgg ctg gcc cag cat cgc gtg ctg 8160  
 Arg Leu Ser Val Gly Thr His Pro Trp Leu Ala Gln His Arg Val Leu  
 2705 2710 2715 2720

ggc gag gtc gtc gtc ccc ggc acc gcc atc ctg gag atg gcc ctg cac 8208  
 Gly Glu Val Val Val Pro Gly Thr Ala Ile Leu Glu Met Ala Leu His  
 2725 2730 2735

gca ggg gcg cgt ctc ggc tgt gac cgg gtg gaa gag ctc acc ctg gaa 8256  
 Ala Gly Ala Arg Leu Gly Cys Asp Arg Val Glu Glu Leu Thr Leu Glu  
 2740 2745 2750

aca ccg ctg gtg gtc ccc gaa cgc gcg gcg ggt gcc ggt agt cgt ggc 8304  
 Thr Pro Leu Val Val Pro Glu Arg Ala Ala Gly Ala Gly Ser Arg Gly  
 2755 2760 2765

cct gcg gga ggg acc aca gtt tca att gaa act gcg gaa gaa cgt gtg 8352  
 Pro Ala Gly Gly Thr Thr Val Ser Ile Glu Thr Ala Glu Glu Arg Val  
 2770 2775 2780

cgg acg aac gac gcc atc gaa atc cag ctg ctg gtg aac gca ccc gac 8400  
 Arg Thr Asn Asp Ala Ile Glu Ile Gln Leu Leu Val Asn Ala Pro Asp



2785	2790	2795	2800	
gaa ggc ggt cgg cga agg gtg tcg ctg tat tcc cgc ccg gcc ggt ggg				8448
Glu Gly Gly Arg Arg Arg Val Ser Leu Tyr Ser Arg Pro Ala Gly Gly				
2805	2810	2815		
tcg aga ggt ggg ggt tgg acg cgc cac gcc acc ggc gaa ctc gtc gtc				8496
Ser Arg Gly Gly Gly Trp Thr Arg His Ala Thr Gly Glu Leu Val Val				
2820	2825	2830		
ggc acc acc ggt ggt agg gcg gtt cct gat tgg tcg gct gag ggt gcc				8544
Gly Thr Thr Gly Gly Arg Ala Val Pro Asp Trp Ser Ala Glu Gly Ala				
2835	2840	2845		
gag tcg att gct ctc gat gag ttc tac gtc gct ctg gcc gga aac ggg				8592
Glu Ser Ile Ala Leu Asp Glu Phe Tyr Val Ala Leu Ala Gly Asn Gly				
2850	2855	2860		
ttc gag tac ggg ccg ttg ttc cag ggg ctt cag gcg gca tgg cgt cgt				8640
Phe Glu Tyr Gly Pro Leu Phe Gln Gly Leu Gln Ala Ala Trp Arg Arg				
2865	2870	2875	2880	
ggt gac gag gtt ctc gcc gaa atc gcc ccg ccg gcc gag gcc gat gcg				8688
Gly Asp Glu Val Leu Ala Glu Ile Ala Pro Pro Ala Glu Ala Asp Ala				
2885	2890	2895		
atg gcg tcg gga tac ctg ctc gac cca gcg ttg ctg gat gcc gcg ctg				8736
Met Ala Ser Gly Tyr Leu Leu Asp Pro Ala Leu Leu Asp Ala Ala Leu				
2900	2905	2910		
cag gcg tcc gcg ctc ggc gac cgc ccg gag caa ggc ggc gcg tgg ctg				8784
Gln Ala Ser Ala Leu Gly Asp Arg Pro Glu Gln Gly Gly Ala Trp Leu				
2915	2920	2925		
ccg ttc tca ttc acc ggc gtc gaa ctt tcc gct ccg gca ggg acg atc				8832
Pro Phe Ser Phe Thr Gly Val Glu Leu Ser Ala Pro Ala Gly Thr Ile				
2930	2935	2940		
agc agg gtg cgg ctg gag acc agg cga ccc gac gcg ata tcg gtg gcc				8880
Ser Arg Val Arg Leu Glu Thr Arg Arg Pro Asp Ala Ile Ser Val Ala				
2945	2950	2955	2960	
gtg atg gat gag agt ggg cgg ttg ctc gcc tcg atc gat tct ctc agg				8928
Val Met Asp Glu Ser Gly Arg Leu Leu Ala Ser Ile Asp Ser Leu Arg				
2965	2970	2975		
cta cga agc gtg tcg tcg gga cag ctg gcg aat cgg gac gct gtc cgc				8976
Leu Arg Ser Val Ser Ser Gly Gln Leu Ala Asn Arg Asp Ala Val Arg				

2980	2985	2990	
gac gcg ctg ttc gag gtg acc tgg gag ccg gtg gcg acg cag tcg acg Asp Ala Leu Phe Glu Val Thr Trp Glu Pro Val Ala Thr Gln Ser Thr 2995 3000 3005			9024
gaa ccg ggt cgc tgg gcc ctg ctt ggt gat act gcc tgc ggt aaa gac Glu Pro Gly Arg Trp Ala Leu Leu Gly Asp Thr Ala Cys Gly Lys Asp 3010 3015 3020			9072
gat ctc atc aaa ctc gca acg gat tcc gcc gac cgc tgc gcg gat ctg Asp Leu Ile Lys Leu Ala Thr Asp Ser Ala Asp Arg Cys Ala Asp Leu 3025 3030 3035 3040			9120
gcg gcg cta gcc gag aaa ctt gat tcc agc gcg ctg gtt cct gat gtc Ala Ala Leu Ala Glu Lys Leu Asp Ser Ser Ala Leu Val Pro Asp Val 3045 3050 3055			9168
gtg gtc tac tgc gcc gga gaa cag gcg gat ccc gcc acc gcc gca gcc Val Val Tyr Cys Ala Gly Glu Gln Ala Asp Pro Gly Thr Gly Ala Ala 3060 3065 3070			9216
gca ctt gcg gag acc cag cag acg ttg gct ctg ctc caa gcg tgg ttg Ala Leu Ala Glu Thr Gln Gln Thr Leu Ala Leu Leu Gln Ala Trp Leu 3075 3080 3085			9264
gct gag ccg cgg ttg gcc gag gca cgt ctg gtg gtg gtg acg tgt gca Ala Glu Pro Arg Leu Ala Glu Ala Arg Leu Val Val Val Thr Cys Ala 3090 3095 3100			9312
gcg gtg acg acg gct ccg agt gac ggt gca tca gag ctg gca cat gcg Ala Val Thr Thr Ala Pro Ser Asp Gly Ala Ser Glu Leu Ala His Ala 3105 3110 3115 3120			9360
ccg ttg tgg ggg ttg ttg cgt gcc gcg cag gtg gag aac ccg ggg cag Pro Leu Trp Gly Leu Leu Arg Ala Ala Gln Val Glu Asn Pro Gly Gln 3125 3130 3135			9408
ttt gtg ctg gcg gac gtc gac gga acc gcc gaa tcg tgg cgt gcg ttg Phe Val Leu Ala Asp Val Asp Gly Thr Ala Glu Ser Trp Arg Ala Leu 3140 3145 3150			9456
ccg agt gcg ttg ggc tcg atg gaa ccg cag ttg gcc ctg cgg aag gcc Pro Ser Ala Leu Gly Ser Met Glu Pro Gln Leu Ala Leu Arg Lys Gly 3155 3160 3165			9504
gcg gtg cga gcg ccc cgc ttg gct tcg gtc gcc ggg cag atc gac gtg Ala Val Arg Ala Pro Arg Leu Ala Ser Val Ala Gly Gln Ile Asp Val			9552

3170	3175	3180	
ccc gcg gtt gtg gcg gat	ccc gac cga acc gtg ctg att tcg ggc ggc	9600	
Pro Ala Val Val Ala Asp	Pro Asp Arg Thr Val Leu Ile Ser Gly Gly		
3185	3190 3195 3200		
acg ggc ctg ttg ggg ggc gcg gtt gcc cgc cac ctg gtg acc gaa cgc	9648		
Thr Gly Leu Leu Gly Gly Ala Val Ala Arg His Leu Val Thr Glu Arg			
3205 3210 3215			
ggt gtc cgc cga ttg gtg ttg acg ggc cgt cgt ggc tgg gat gct cct	9696		
Gly Val Arg Arg Leu Val Leu Thr Gly Arg Arg Gly Trp Asp Ala Pro			
3220 3225 3230			
gga atc acc gag ttg gtg ggt gag ctg aac ggc ctc ggt gcc gtg gtc	9744		
Gly Ile Thr Glu Leu Val Gly Glu Leu Asn Gly Leu Gly Ala Val Val			
3235 3240 3245			
gac gtg gtg gcg tgc gac gtc gcg gat cgt gct gat ctg gag tcg ttg	9792		
Asp Val Val Ala Cys Asp Val Ala Asp Arg Ala Asp Leu Glu Ser Leu			
3250 3255 3260			
ctg gcg gcg gtc ccg gcg gaa ttt ccg ttg tgc ggc gtg gtg cat gcc	9840		
Leu Ala Ala Val Pro Ala Glu Phe Pro Leu Cys Gly Val Val His Ala			
3265 3270 3275 3280			
gcg ggg gcg ctg gcc gac ggg gtg atc gag tcg ttg tca ccg gac gac	9888		
Ala Gly Ala Leu Ala Asp Gly Val Ile Glu Ser Leu Ser Pro Asp Asp			
3285 3290 3295			
gtg gga gcg gtg ttc ggc ccg aag gcg gcg ggg gcg tgg aat ctg cac	9936		
Val Gly Ala Val Phe Gly Pro Lys Ala Ala Gly Ala Trp Asn Leu His			
3300 3305 3310			
gag ctg act cgt gat acg gac ctg tcg ttc ttc gcg ttg ttc tcc tcg	9984		
Glu Leu Thr Arg Asp Thr Asp Leu Ser Phe Phe Ala Leu Phe Ser Ser			
3315 3320 3325			
ctt tcc ggt gtt gcc ggc gct cct ggt cag ggc aat tat gcg gcg gcg	10032		
Leu Ser Gly Val Ala Gly Ala Pro Gly Gln Gly Asn Tyr Ala Ala Ala			
3330 3335 3340			
aac gcg ttc ctg gac gca ttg gcg cat tac cgg cgg tca cag gga ctg	10080		
Asn Ala Phe Leu Asp Ala Leu Ala His Tyr Arg Arg Ser Gln Gly Leu			
3345 3350 3355 3360			
cct gcg gtg tcg ctg gcc tgg ggc ctg tgg gag cag ccg agc ggg atg	10128		
Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro Ser Gly Met			

3365	3370	3375	
acg gag acg ctc agc gag gtc gac cgg agc agg atc gcg cgc gcc aac			10176
Thr Glu Thr Leu Ser Glu Val Asp Arg Ser Arg Ile Ala Arg Ala Asn			
3380	3385	3390	
ccg ccg ttg tcc acc aag gag gga ttg cgg ctg ttc gat gcc ggg ctg			10224
Pro Pro Leu Ser Thr Lys Glu Gly Leu Arg Leu Phe Asp Ala Gly Leu			
3395	3400	3405	
gcg ctg gac cgg gca gcg gta gtt ccg gcg aag ttg gac agg act ttc			10272
Ala Leu Asp Arg Ala Ala Val Val Pro Ala Lys Leu Asp Arg Thr Phe			
3410	3415	3420	
ctg gcc gag cag gcg cgg tcg ggc tcg ctg ccc gca ttg ttg acg gca			10320
Leu Ala Glu Gln Ala Arg Ser Gly Ser Leu Pro Ala Leu Leu Thr Ala			
3425	3430	3435	3440
ctg gta ccc ccc atc cgt cgt aat agg cgg gct agc gga acc gag ctc			10368
Leu Val Pro Pro Ile Arg Arg Asn Arg Arg Ala Ser Gly Thr Glu Leu			
3445	3450	3455	
gcg gac gag ggc acc ctg ctc ggg gtg gtg cgg gag cat gcc gcg gcc			10416
Ala Asp Glu Gly Thr Leu Leu Gly Val Val Arg Glu His Ala Ala Ala			
3460	3465	3470	
gtg ctg ggg tat tcg agc gcg gct gac gtc ggg gtc gag cgc gct ttc			10464
Val Leu Gly Tyr Ser Ser Ala Ala Asp Val Gly Val Glu Arg Ala Phe			
3475	3480	3485	
cgg gat ctg ggt ttt gat tcg ttg tct ggt gtg gag ttg cgg aac cgc			10512
Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg Asn Arg			
3490	3495	3500	
ctt gcc ggg gtg ctg ggg gtg cgg ttg ccg gcg act gcg gtg ttc gac			10560
Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val Phe Asp			
3505	3510	3515	3520
tat ccg acg ccg agg gcg ctg gcc cgg ttc ctg cac cag gaa ctg gca			10608
Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu Leu Ala			
3525	3530	3535	
gac gag atc gct acg acg cca gcg ccg gtg acg acg acc agg gca ccg			10656
Asp Glu Ile Ala Thr Thr Pro Ala Pro Val Thr Thr Thr Arg Ala Pro			
3540	3545	3550	
gtc gcc gaa gac gat ctc gtc gcg ata gtc ggg atg gga tgc cgt ttt			10704
Val Ala Glu Asp Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe			

3555	3560	3565	
ccc ggt cag gtg tcc tcg ccg gag gag ctc tgg cgt ttg gtg gcc ggg			10752
Pro Gly Gln Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly			
3570	3575	3580	
ggc gtg gat gcg gtc gcg gac ttc cca gcc gat cgc ggc tgg gat ctg			10800
Gly Val Asp Ala Val Ala Asp Phe Pro Ala Asp Arg Gly Trp Asp Leu			
3585	3590	3595	3600
gca ggc ttg ttc gat ccg gac ccg gaa cgg gct ggg aag acc tac gtg			10848
Ala Gly Leu Phe Asp Pro Asp Pro Glu Arg Ala Gly Lys Thr Tyr Val			
3605	3610	3615	
cgg gaa ggg gcc ttc ctc acc gac gcc gat cgg ttc gat gcg ggt ttc			10896
Arg Glu Gly Ala Phe Leu Thr Asp Ala Asp Arg Phe Asp Ala Gly Phe			
3620	3625	3630	
ttc ggg att tcc ccg cgt gag gcg ttg gcg atg gat ccg cag caa cgg			10944
Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg			
3635	3640	3645	
ctg ttg ctg gag ctg tcc tgg gag gcc att gaa cgg gca ggg atc gat			10992
Leu Leu Leu Glu Leu Ser Trp Glu Ala Ile Glu Arg Ala Gly Ile Asp			
3650	3655	3660	
ccg ggt tcg ctg agg ggg agt cgg acc ggt gtg ttc gcg ggg ctg atg			11040
Pro Gly Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met			
3665	3670	3675	3680
tac cac gac tat ggc gcc cgg ttc gcc agc cga gcc ccg gaa ggt ttc			11088
Tyr His Asp Tyr Gly Ala Arg Phe Ala Ser Arg Ala Pro Glu Gly Phe			
3685	3690	3695	
gag ggg tat ctc ggc aat ggc agt gct ggg agt gtc gcg tcg ggc cgg			11136
Glu Gly Tyr Leu Gly Asn Gly Ser Ala Gly Ser Val Ala Ser Gly Arg			
3700	3705	3710	
att gcg tac tcg ttt ggt ttc gag ggt cct gcg gtg acg gtg gat act			11184
Ile Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr			
3715	3720	3725	
gcg tgt tcg tcg tcg ttg gtg gcg ttg cat ttg gcg ggt cag tcg ttg			11232
Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ser Leu			
3730	3735	3740	
cgt tcc ggc gaa tgc gat ctc gcc ctt gcc ggt ggt gtg acg gtg atg			11280
Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met			

3745	3750	3755	3760	
tcg acg ccc ggg acg ttt gtg gaa ttc tcc cgt cag cgg ggc ctg gca				11328
Ser Thr Pro Gly Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala				
3765	3770	3775		
ccg gac ggg cgg tgc aag tcg ttc gcg gag agc gcg gac ggt acc ggt				11376
Pro Asp Gly Arg Cys Lys Ser Phe Ala Glu Ser Ala Asp Gly Thr Gly				
3780	3785	3790		
tgg ggt gag ggt gct ggt ttg gtg ttg ttg gag cgg ttg tcg gat gct				11424
Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala				
3795	3800	3805		
cgg cgg aat ggg cat cgg gtg ttg gcg gtg gtt cgt ggg tcg gcg gtg				11472
Arg Arg Asn Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val				
3810	3815	3820		
aat cag gat ggt gcg tcg aat ggc ttg acc gcg ccg aat ggt ccc tcg				11520
Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser				
3825	3830	3835	3840	
cag cag cgg gtc atc cag cag gcg ttg gcg agt gcg ggt ctg tcg gtg				11568
Gln Gln Arg Val Ile Gln Gln Ala Leu Ala Ser Ala Gly Leu Ser Val				
3845	3850	3855		
tcc gat gtg gat gcc gtg gag gcg cat ggg acc ggg acc agg ttg ggt				11616
Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly				
3860	3865	3870		
gat ccg att gag gcg cag gct ctg att gct acg tat ggg cgc gat cgt				11664
Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg				
3875	3880	3885		
gat ccc ggt cgg ccg ttg tgg ttg ggg tcg gtg aag tcc aac atc ggt				11712
Asp Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly				
3890	3895	3900		
cat acg cag gcg gcg gcg ggt gtt gcc ggt gtg atc aag atg gtg atg				11760
His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met				
3905	3910	3915	3920	
gcg atg cgg cac ggg caa ctt ccg cgc acg ctg cac gtg gat gca ccc				11808
Ala Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Asp Ala Pro				
3925	3930	3935		
tcc tcg cag gtg gat tgg tcg gcg ggg agg gtc cag ctc ctg acg gag				11856
Ser Ser Gln Val Asp Trp Ser Ala Gly Arg Val Gln Leu Leu Thr Glu				

3940	3945	3950	
aac acg ccc tgg ccc gac agt ggt cgc ccc tgt cgg gtg ggg gtg tcg			11904
Asn Thr Pro Trp Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val Ser			
3955	3960	3965	
tcg ttc ggg atc agc ggc acc aac gcg cac gtc atc ctg gaa cag tcc			11952
Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ser			
3970	3975	3980	
acg ggg cag atg gat cag gca gcg gag ccg gat tcg agt cct gtt ctg			12000
Thr Gly Gln Met Asp Gln Ala Ala Glu Pro Asp Ser Ser Pro Val Leu			
3985	3990	3995	4000
gat gtt ccg gtg gtg ccg tgg gtg gtg tcg ggc aaa aca ccc gaa gcg			12048
Asp Val Pro Val Val Pro Trp Val Val Ser Gly Lys Thr Pro Glu Ala			
4005	4010	4015	
cta tcc gcc cag gcg gca acg ttg gcg acc tat ttg gac caa aat gtt			12096
Leu Ser Ala Gln Ala Ala Thr Leu Ala Thr Tyr Leu Asp Gln Asn Val			
4020	4025	4030	
gat gtc tcc cct ctg gac gtt ggg att tcg ctt gcg gtg acc cgt tcg			12144
Asp Val Ser Pro Leu Asp Val Gly Ile Ser Leu Ala Val Thr Arg Ser			
4035	4040	4045	
gcg ctg gat gag ccg gcg gtg gtg ctg ggg tcg gat cgt gac acg ttg			12192
Ala Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Asp Thr Leu			
4050	4055	4060	
ttg tct ggc ctg aat gcg ctg gct gcc ggt cat gag gct gct ggc gtg			12240
Leu Ser Gly Leu Asn Ala Leu Ala Ala Gly His Glu Ala Ala Gly Val			
4065	4070	4075	4080
gtt acg gga cct gtc ggg att ggt ggc ccg acc ggg ttt gtg ttc gcc			12288
Val Thr Gly Pro Val Gly Ile Gly Gly Arg Thr Gly Phe Val Phe Ala			
4085	4090	4095	
ggt caa ggc ggt cag tgg ttg ggg atg ggc cgc ccg ttg tac tcg gag			12336
Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Arg Leu Tyr Ser Glu			
4100	4105	4110	
ttt ccg gcg ttc gcc ggt gct ttc gac gaa gca tgc gcc gag ctc gat			12384
Phe Pro Ala Phe Ala Gly Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp			
4115	4120	4125	
gcg aac ctg ggg agg gaa gtc ggg gtt ccg gat gtg gtg ttc ggc tcc			12432
Ala Asn Leu Gly Arg Glu Val Gly Val Arg Asp Val Val Phe Gly Ser			

4130	4135	4140	
gac gag tcc ttg ctg gat cgg act ttg tgg gcg cag tcg ggt ttg ttc			12480
Asp Glu Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Phe			
4145	4150	4155	4160
gcg ttg cag gtc ggt ctc tgg gaa ttg ttg ggt acg tgg ggt gtt cgg			12528
Ala Leu Gln Val Gly Leu Trp Glu Leu Leu Gly Thr Trp Gly Val Arg			
4165	4170	4175	
ccc agc gta gtg ctg ggg cat tcg gtc ggg gag cta gcc gcg gcg ttc			12576
Pro Ser Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe			
4180	4185	4190	
gcc gca ggt gtg ctg tcg atg gcg gag gcg gct cgg ctg gtg gcg ggt			12624
Ala Ala Gly Val Leu Ser Met Ala Glu Ala Ala Arg Leu Val Ala Gly			
4195	4200	4205	
cgt gcg cgg ttg atg cag gcg ttg cct tct ggc ggt gcc atg ctg gcg			12672
Arg Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala Met Leu Ala			
4210	4215	4220	
gtg tcc gcg acc gag gcc cga gtc ggc ccg ctg ctc gat ggg gtg cgg			12720
Val Ser Ala Thr Glu Ala Arg Val Gly Pro Leu Leu Asp Gly Val Arg			
4225	4230	4235	4240
gat cgt gtt ggt gtc gca gcg gtt aac gct ccg ggg tcg gtg gtg ctt			12768
Asp Arg Val Gly Val Ala Ala Val Asn Ala Pro Gly Ser Val Val Leu			
4245	4250	4255	
tcc ggt gac cgg gat gtg ctc gat ggc att gcc ggt cgg ctg gac ggg			12816
Ser Gly Asp Arg Asp Val Leu Asp Gly Ile Ala Gly Arg Leu Asp Gly			
4260	4265	4270	
caa ggt atc cgg tcg agg tgg ttg cgg gtt tcg cac gcg ttt cat tcg			12864
Gln Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser			
4275	4280	4285	
cat cgg atg gat ccg atg ctg gcg gag ttc gcc gag ctc gca cgg agc			12912
His Arg Met Asp Pro Met Leu Ala Glu Phe Ala Glu Leu Ala Arg Ser			
4290	4295	4300	
gtg gac tac cgg tct cca cgg ctg ccg att gtc tcg acg ctg acc gga			12960
Val Asp Tyr Arg Ser Pro Arg Leu Pro Ile Val Ser Thr Leu Thr Gly			
4305	4310	4315	4320
aac ctc gat gac gtg ggc gtg atg gct acg ccg gag tat tgg gtg cgc			13008
Asn Leu Asp Asp Val Gly Val Met Ala Thr Pro Glu Tyr Trp Val Arg			



4325	4330	4335	
cag gtg cga gag ccc gtc cgc ttc gcc gac ggt gtc cag gcg ctt gtg			13056
Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Gln Ala Leu Val			
4340	4345	4350	
gac caa ggc gtc gac acg att gtg gaa ctc ggt ccg gac ggg gcg ttg			13104
Asp Gln Gly Val Asp Thr Ile Val Glu Leu Gly Pro Asp Gly Ala Leu			
4355	4360	4365	
tcg agc ttg gtt caa gag tgt gtg gcg gag tcc ggg cgg gcg acg ggg			13152
Ser Ser Leu Val Gln Glu Cys Val Ala Glu Ser Gly Arg Ala Thr Gly			
4370	4375	4380	
att ccg ttg gtg cgg aga gac cgt gat gag gtc cga acg gtg ctg gac			13200
Ile Pro Leu Val Arg Arg Asp Arg Asp Glu Val Arg Thr Val Leu Asp			
4385	4390	4395	4400
gct ttg gcg cag acc cac act cgt ggt ggc gcg gtg gac tgg ggg tca			13248
Ala Leu Ala Gln Thr His Thr Arg Gly Gly Ala Val Asp Trp Gly Ser			
4405	4410	4415	
ttt ttc gct ggt acg agg gca acg caa gtc gac ctt ccc acg tat gcc			13296
Phe Phe Ala Gly Thr Arg Ala Thr Gln Val Asp Leu Pro Thr Tyr Ala			
4420	4425	4430	
ttc caa cga cag cgg tac tgg ctg gag cca tcg gat tcc ggt gat gtg			13344
Phe Gln Arg Gln Arg Tyr Trp Leu Glu Pro Ser Asp Ser Gly Asp Val			
4435	4440	4445	
acc ggt gtt ggc ctg acc ggg gcg gag cat ccg ctg ttg ggt gcc gtg			13392
Thr Gly Val Gly Leu Thr Gly Ala Glu His Pro Leu Leu Gly Ala Val			
4450	4455	4460	
gtg ccg gtc gcg ggc ggc gat gag gtg ctg ctg acc ggc agg ctg tcg			13440
Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg Leu Ser			
4465	4470	4475	4480
gtg ggg acg cat ccg tgg ctg gcg gaa cac cgc gtg ctg ggc gaa gtc			13488
Val Gly Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly Glu Val			
4485	4490	4495	
gtc gtc ccc ggc acc gcg ttg ctg gag atg gcg tgg cgg gcc ggt agc			13536
Val Val Pro Gly Thr Ala Leu Leu Glu Met Ala Trp Arg Ala Gly Ser			
4500	4505	4510	
cag gtc ggt tgt gaa cgt gtg gag gag ctc acc ttg gag gca ccg ctg			13584
Gln Val Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala Pro Leu			

4515	4520	4525	
gtc ctg ccg gag cgg ggc gct gcg gcg gtg cag ttg gcg gtg ggg gct Val Leu Pro Glu Arg Gly Ala Ala Ala Val Gln Leu Ala Val Gly Ala 4530 4535 4540			13632
ccg gat gag gcc ggc cgg cgc agt ttg cag ctc tat tcc cga ggc gct Pro Asp Glu Ala Gly Arg Arg Ser Leu Gln Leu Tyr Ser Arg Gly Ala 4545 4550 4555 4560			13680
gat gaa gac ggc gac tgg cgg cgg att gcc tcc ggg ctg ttg gcc cag Asp Glu Asp Gly Asp Trp Arg Arg Ile Ala Ser Gly Leu Leu Ala Gln 4565 4570 4575			13728
gcc aat gcg gtg ccg ccg gcg gat tcg acg gca tgg ccg ccg gac ggc Ala Asn Ala Val Pro Pro Ala Asp Ser Thr Ala Trp Pro Pro Asp Gly 4580 4585 4590			13776
gcc ggg cag gtc gat ctg gcg gag ttc tac gag cgc ctc gcc gag cgc Ala Gly Gln Val Asp Leu Ala Glu Phe Tyr Glu Arg Leu Ala Glu Arg 4595 4600 4605			13824
ggc ttg acc tac ggt ccg gta ttc caa ggg ctc cgc gcc gca tgg cgg Gly Leu Thr Tyr Gly Pro Val Phe Gln Gly Leu Arg Ala Ala Trp Arg 4610 4615 4620			13872
cac ggc gac gat atc ttc gcc gaa ttg gcc ggg tca cca gac gcc tcg His Gly Asp Asp Ile Phe Ala Glu Leu Ala Gly Ser Pro Asp Ala Ser 4625 4630 4635 4640			13920
ggt ttc ggc atc cac ccg gcg ctg ctg gac gct gca ctg cac gcg atg Gly Phe Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu His Ala Met 4645 4650 4655			13968
gcg ctt ggt gct tcg ccc gac tcg gaa gcg cgt ctg ccg ttt tcc tgg Ala Leu Gly Ala Ser Pro Asp Ser Glu Ala Arg Leu Pro Phe Ser Trp 4660 4665 4670			14016
cgt ggc gcc cag ctg tac cgc gct gaa gga gca gcg ctt cgg gta cgg Arg Gly Ala Gln Leu Tyr Arg Ala Glu Gly Ala Ala Leu Arg Val Arg 4675 4680 4685			14064
ctc tcg ccg ctg ggc tcc ggt gca gtc tca ttg acg ttg gtg gat gcc Leu Ser Pro Leu Gly Ser Gly Ala Val Ser Leu Thr Leu Val Asp Ala 4690 4695 4700			14112
aca ggg cga cga gtc gct gcg gtg gaa tcg ctt tcg acg cga ccg gtc Thr Gly Arg Arg Val Ala Ala Val Glu Ser Leu Ser Thr Arg Pro Val			14160

4705	4710	4715	4720	
tcc acc gac cag atc ggt gcc ggt cgc ggc gat caa gag cgg ctg ctg				14208
Ser Thr Asp Gln Ile Gly Ala Gly Arg Gly Asp Gln Glu Arg Leu Leu				
4725	4730	4735		
cac gtc gag tgg gta agg tcg gct gaa tct gcg ggg atg tct ctg acc				14256
His Val Glu Trp Val Arg Ser Ala Glu Ser Ala Gly Met Ser Leu Thr				
4740	4745	4750		
tcc tgc gcg gtg gtc ggt ttg ggc gaa ccg gag tgg cac gct gcg ctg				14304
Ser Cys Ala Val Val Gly Leu Gly Glu Pro Glu Trp His Ala Ala Leu				
4755	4760	4765		
aag acc act ggt gtc caa gtc gag tcc cat gcg gac ctt gct tcg ttg				14352
Lys Thr Thr Gly Val Gln Val Glu Ser His Ala Asp Leu Ala Ser Leu				
4770	4775	4780		
gcc acc gag gtt gcc aag cgg ggt tca gct cct ggt gcg gtc atc gtc				14400
Ala Thr Glu Val Ala Lys Arg Gly Ser Ala Pro Gly Ala Val Ile Val				
4785	4790	4795	4800	
ccg tgc ccg cga ccc cga gcg atg cag gag ctg ccg acc gcc gcg cga				14448
Pro Cys Pro Arg Pro Arg Ala Met Gln Glu Leu Pro Thr Ala Ala Arg				
4805	4810	4815		
agg gcg acg caa cag gcg atg gcg atg ctg cag caa tgg ctt gcc gat				14496
Arg Ala Thr Gln Gln Ala Met Ala Met Leu Gln Gln Trp Leu Ala Asp				
4820	4825	4830		
gac cgg ttc gtc agt acg cgc ctg atc ctg ctg acg cat cgg gcg gtc				14544
Asp Arg Phe Val Ser Thr Arg Leu Ile Leu Leu Thr His Arg Ala Val				
4835	4840	4845		
tcc gca gtt gct gga gaa gac gtg ctc gac ctg gta cac gcg ccg ctg				14592
Ser Ala Val Ala Gly Glu Asp Val Leu Asp Leu Val His Ala Pro Leu				
4850	4855	4860		
tgg ggc ttg gtc cgc agc gcg caa gcg gag cac ccg gac cga ttc gcc				14640
Trp Gly Leu Val Arg Ser Ala Gln Ala Glu His Pro Asp Arg Phe Ala				
4865	4870	4875	4880	
ttg atc gat atg gac gac gag cga gca tcg cag acg gca ctc gcc gaa				14688
Leu Ile Asp Met Asp Asp Glu Arg Ala Ser Gln Thr Ala Leu Ala Glu				
4885	4890	4895		
gcg ctg act gcg gga gaa gcg cag ctc gcg gtg cgg tcg gga gtt gtg				14736
Ala Leu Thr Ala Gly Glu Ala Gln Leu Ala Val Arg Ser Gly Val Val				

4900	4905	4910	
ctg gcg ccc cgc ctc ggc cag gtg aag gtg agt gga ggt gaa gcg ttc Leu Ala Pro Arg Leu Gly Gln Val Lys Val Ser Gly Gly Glu Ala Phe 4915 4920 4925			14784
agg tgg gat gaa ggc acc gtg ctg gtc acc ggc gga acc ggc ggg ctc Arg Trp Asp Glu Gly Thr Val Leu Val Thr Gly Gly Thr Gly Gly Leu 4930 4935 4940			14832
ggg gcc ctg ctc gca cgc cat ctg gtc agc gcc cac ggt gtg cgg cac Gly Ala Leu Leu Ala Arg His Leu Val Ser Ala His Gly Val Arg His 4945 4950 4955 4960			14880
ctg ttg ctc gca agt cgc cgt ggt ctg gcg gcg ccc gga gcg gat gag Leu Leu Leu Ala Ser Arg Arg Gly Leu Ala Ala Pro Gly Ala Asp Glu 4965 4970 4975			14928
ctg gtg gcc gag ctg gag cag gcc ggc gcc gac gtc gcg gtc gtc gcg Leu Val Ala Glu Leu Glu Gln Ala Gly Ala Asp Val Ala Val Val Ala 4980 4985 4990			14976
tgc gac tcg gca gat cgg gac tcg ctt gcg cgg ctg gtg gcg tcg gtg Cys Asp Ser Ala Asp Arg Asp Ser Leu Ala Arg Leu Val Ala Ser Val 4995 5000 5005			15024
cct gcg gaa aac ccg ttg cgg gtg gtg gtg cac gcc gcc ggt gtg ctg Pro Ala Glu Asn Pro Leu Arg Val Val Val His Ala Ala Gly Val Leu 5010 5015 5020			15072
gat gac ggt gtg ctg atg tcg atg tcg ccg gag cgc ttg gac gcg gtg Asp Asp Gly Val Leu Met Ser Met Ser Pro Glu Arg Leu Asp Ala Val 5025 5030 5035 5040			15120
ttg cgg ccc aaa gtg gat gcc gcg tgg tac ctg cac gag ctg act cgg Leu Arg Pro Lys Val Asp Ala Ala Trp Tyr Leu His Glu Leu Thr Arg 5045 5050 5055			15168
gaa ctc ggt ctg tcg gcg ttc gtg ttg ttc tcc tcg gtc gcg ggc ctg Glu Leu Gly Leu Ser Ala Phe Val Leu Phe Ser Ser Val Ala Gly Leu 5060 5065 5070			15216
ttc ggc ggt gcg ggg cag agc aat tac gct gcc ggc aac gct ttc ctg Phe Gly Gly Ala Gly Gln Ser Asn Tyr Ala Ala Gly Asn Ala Phe Leu 5075 5080 5085			15264
gat gcc ttg gcg cat tgc cgg cag gcc cag ggg ctg ccc gcg ctg tcg Asp Ala Leu Ala His Cys Arg Gln Ala Gln Gly Leu Pro Ala Leu Ser			15312

5090	5095	5100	
ctg gcc tcc ggg ctg tgg gcg agt atc gat gga atg gcg ggc gac ctc			15360
Leu Ala Ser Gly Leu Trp Ala Ser Ile Asp Gly Met Ala Gly Asp Leu			
5105	5110	5115	5120
gct gcg gca gat gtg gag cgg ctg tgg cgg gca ggc att ggc ccg ctt			15408
Ala Ala Ala Asp Val Glu Arg Leu Ser Arg Ala Gly Ile Gly Pro Leu			
5125	5130	5135	
tcg gca ccg gga ggg ctg gcc ttg ttc gac gct gcc gtt ggc tcg gac			15456
Ser Ala Pro Gly Gly Leu Ala Leu Phe Asp Ala Ala Val Gly Ser Asp			
5140	5145	5150	
gaa ccg ttg ctg gca ccg gtg cga ctg gat gtc gaa gca ctg cgt gtg			15504
Glu Pro Leu Leu Ala Pro Val Arg Leu Asp Val Glu Ala Leu Arg Val			
5155	5160	5165	
cag gcc cga tcc gtg cag acc cgg att ccg gaa atg ctg cat ggc atg			15552
Gln Ala Arg Ser Val Gln Thr Arg Ile Pro Glu Met Leu His Gly Met			
5170	5175	5180	
gca atg ggg cca agc cgc cgc act ccg ttc act tcc agg gtt gag ccg			15600
Ala Met Gly Pro Ser Arg Arg Thr Pro Phe Thr Ser Arg Val Glu Pro			
5185	5190	5195	5200
ttg cac gaa cgg ctg gcc gga ttg tcg gag ggc gaa cgt cgg cag caa			15648
Leu His Glu Arg Leu Ala Gly Leu Ser Glu Gly Glu Arg Arg Gln Gln			
5205	5210	5215	
gtg ctc cag cgc gtc cgc gcc gat atc gcg gtg gta ctg ggg cac ggc			15696
Val Leu Gln Arg Val Arg Ala Asp Ile Ala Val Val Leu Gly His Gly			
5220	5225	5230	
agg tcg agc gat gtg gac atc gag aag cct ttg gcc gag ctg ggt ttc			15744
Arg Ser Ser Asp Val Asp Ile Glu Lys Pro Leu Ala Glu Leu Gly Phe			
5235	5240	5245	
gac tcg ctg acg gcc atc gaa ctc cgc aac cgt ctc gct acc gcc acc			15792
Asp Ser Leu Thr Ala Ile Glu Leu Arg Asn Arg Leu Ala Thr Ala Thr			
5250	5255	5260	
gga ctg cgg ctt ccc gcg acg ctg gcc ttc gac cac ggc act gcg gcg			15840
Gly Leu Arg Leu Pro Ala Thr Leu Ala Phe Asp His Gly Thr Ala Ala			
5265	5270	5275	5280
gca ctc gcc cag cac gtg tgc gcg cag cta ggc acc gcg acc gcg ccg			15888
Ala Leu Ala Gln His Val Cys Ala Gln Leu Gly Thr Ala Thr Ala Pro			

5285	5290	5295	
gca ccg agg cga acc gac gac aac gac gcc acg gag ccc gtg agg tcg Ala Pro Arg Arg Thr Asp Asp Asn Asp Ala Thr Glu Pro Val Arg Ser			15936
5300	5305	5310	
ctc ttc caa cag gcg tat gcg gct ggc cgg ata ctt gac ggg atg gat Leu Phe Gln Gln Ala Tyr Ala Ala Gly Arg Ile Leu Asp Gly Met Asp			15984
5315	5320	5325	
ttg gtg aag gtc gct gcc cag ttg cga ccg gtg ttc ggt tcg cct ggc Leu Val Lys Val Ala Ala Gln Leu Arg Pro Val Phe Gly Ser Pro Gly			16032
5330	5335	5340	
gag ctg gaa tcc ctg ccg aaa ccc gtc cag ctt tcc cgt ggt ccc gaa Glu Leu Glu Ser Leu Pro Lys Pro Val Gln Leu Ser Arg Gly Pro Glu			16080
5345	5350	5355	5360
gag ctt gcc ttg gtg tgc atg ccg gcg ctg atc ggg atg ccg ccc gca Glu Leu Ala Leu Val Cys Met Pro Ala Leu Ile Gly Met Pro Pro Ala			16128
5365	5370	5375	
cag cag tac gcg cgg atc gcc gcc ggg ttc cgc gat gtg cgg gac gtt Gln Gln Tyr Ala Arg Ile Ala Ala Gly Phe Arg Asp Val Arg Asp Val			16176
5380	5385	5390	
tcg gtg atc ccg atg cct gga ttc att gcg gga gaa ccg ctg ccg tcc Ser Val Ile Pro Met Pro Gly Phe Ile Ala Gly Glu Pro Leu Pro Ser			16224
5395	5400	5405	
gcc atc gag gtg gcg gtt cgg acg cag gcg gag gcg gtg ctg cag gaa Ala Ile Glu Val Ala Val Arg Thr Gln Ala Glu Ala Val Leu Gln Glu			16272
5410	5415	5420	
ttc gcc ggg ggc tcg ttc gta ctg gtc ggg cat tcc tcc ggg ggc tgg Phe Ala Gly Gly Ser Phe Val Leu Val Gly His Ser Ser Gly Gly Trp			16320
5425	5430	5435	5440
ctg gcg cac gag gta gcc ggt gag ctg gag cgt cgc ggg gtc gtc ccg Leu Ala His Glu Val Ala Gly Glu Leu Glu Arg Arg Gly Val Val Pro			16368
5445	5450	5455	
gcc ggg gtc gta ctg ctg gac acc tac atc ccc ggt gag atc acg ccg Ala Gly Val Val Leu Leu Asp Thr Tyr Ile Pro Gly Glu Ile Thr Pro			16416
5460	5465	5470	
agg ttc tcc gtg gcg atg gcc cac cgg acg tat gag aag ctc gcg act Arg Phe Ser Val Ala Met Ala His Arg Thr Tyr Glu Lys Leu Ala Thr			16464

5475

5480

5485

ttc acg gac atg cag gat gtc ggt atc acc gcg atg ggc ggg tac ttc 16512  
 Phe Thr Asp Met Gln Asp Val Gly Ile Thr Ala Met Gly Gly Tyr Phe  
 5490 5495 5500

cgg atg ttc acc gag tgg act ccg acg ccg atc ggt gct ccg acg ctg 16560  
 Arg Met Phe Thr Glu Trp Thr Pro Thr Pro Ile Gly Ala Pro Thr Leu  
 5505 5510 5515 5520

ttc gtg cgg acc gaa gat tgc gtc gca gac cct gaa ggg cgg ccg tgg 16608  
 Phe Val Arg Thr Glu Asp Cys Val Ala Asp Pro Glu Gly Arg Pro Trp  
 5525 5530 5535

aca gat gac tcc tgg cgg cca ggg tgg act ctc gcg gat gcc acg gtc 16656  
 Thr Asp Asp Ser Trp Arg Pro Gly Trp Thr Leu Ala Asp Ala Thr Val  
 5540 5545 5550

cag gtg ccg ggc gac cac ttc tcg atg atg gac gag cac gcc ggg tcc 16704  
 Gln Val Pro Gly Asp His Phe Ser Met Met Asp Glu His Ala Gly Ser  
 5555 5560 5565

acc gca cag gca gtc gcg agt tgg ctt gac aaa ctc aac cag cgc acc 16752  
 Thr Ala Gln Ala Val Ala Ser Trp Leu Asp Lys Leu Asn Gln Arg Thr  
 5570 5575 5580

gct cgg caa cgc tga 16767  
 Ala Arg Gln Arg  
 5585

&lt;210&gt; 50

&lt;211&gt; 5588

&lt;212&gt; PRT

&lt;213&gt; Saccharopolyspora spinosa

&lt;400&gt; 50

Met Ala Asn Glu Glu Lys Leu Arg Glu Tyr Leu Lys Arg Val Val Val  
 1 5 10 15

Glu Leu Glu Glu Ala His Glu Arg Leu His Glu Leu Glu Arg Gln Glu  
 20 25 30

His Asp Pro Ile Ala Ile Val Ser Met Gly Cys Arg Tyr Pro Gly Gly  
 35 40 45

Val Ser Thr Pro Glu Glu Leu Trp Arg Leu Val Val Asp Gly Gly Asp  
 50 55 60

Ala Ile Ala Asn Phe Pro Glu Asp Arg Gly Trp Asn Leu Asp Glu Leu  
 65 70 75 80

Phe Asp Pro Asp Pro Gly Arg Ala Gly Thr Ser Tyr Val Arg Glu Gly  
 85 90 95

Gly Phe Leu Arg Gly Val Ala Asp Phe Asp Ala Gly Leu Phe Gly Ile  
 100 105 110

Ser Pro Arg Glu Ala Gln Ala Met Asp Pro Gln Gln Arg Leu Leu Leu  
 115 120 125

Glu Ile Ser Trp Glu Val Phe Glu Arg Ala Gly Ile Asp Pro Phe Ser  
 130 135 140

Leu Arg Gly Thr Lys Thr Gly Val Phe Ala Gly Leu Ile Tyr His Asp  
 145 150 155 160

Tyr Ala Ser Arg Phe Arg Lys Thr Pro Ala Glu Phe Glu Gly Tyr Phe  
 165 170 175

Ala Thr Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Thr  
 180 185 190

Phe Gly Leu Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser  
 195 200 205

Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu Arg Leu Gly Glu  
 210 215 220

Cys Asp Leu Ala Leu Ala Gly Gly Ile Ser Val Met Ala Thr Pro Gly  
 225 230 235 240

Ala Phe Val Glu Phe Ser Arg Gln Arg Ala Leu Ala Ser Asp Gly Arg  
 245 250 255

Cys Lys Pro Phe Ala Asp Ala Ala Asp Gly Thr Gly Trp Gly Glu Gly  
 260 265 270

Ala Gly Met Leu Leu Leu Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly  
 275 280 285

His Pro Val Leu Ala Ala Val Val Gly Ser Ala Ile Asn Gln Asp Gly  
 290 295 300

Thr Ser Asn Gly Leu Thr Ala Pro Ser Gly Pro Ala Gln Gln Arg Val  
 305 310 315 320



Ile Arg Gln Ala Leu Ala Asn Ala Gly Leu Ser Pro Ala Glu Val Asp  
325 330 335

Val Val Glu Ala His Gly Thr Gly Thr Ala Leu Gly Asp Pro Ile Glu  
340 345 350

Ala Gln Ala Leu Ile Ala Thr Tyr Gly Ala Asn Arg Ser Ala Asp His  
355 360 365

Pro Leu Leu Leu Gly Ser Leu Lys Ser Asn Ile Gly His Thr Gln Ala  
370 375 380

Ala Ala Gly Val Ala Gly Val Ile Lys Ser Val Leu Ala Ile Arg His  
385 390 395 400

Arg Glu Met Pro Arg Ser Leu His Ile Asp Gln Pro Ser Gln His Val  
405 410 415

Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Asp Ser Val Asp Trp  
420 425 430

Pro Asp Leu Gly Arg Pro Arg Arg Ala Gly Val Ser Ser Phe Gly Met  
435 440 445

Ser Gly Thr Asn Ala His Leu Ile Val Glu Glu Val Ser Asp Glu Pro  
450 455 460

Val Ser Gly Ser Thr Glu Pro Thr Gly Ala Phe Pro Trp Pro Leu Ser  
465 470 475 480

Gly Lys Thr Glu Thr Ala Leu Arg Glu Gln Ala Ala Glu Leu Leu Ser  
485 490 495

Val Val Thr Glu His Pro Glu Pro Gly Leu Gly Asp Val Gly Tyr Ser  
500 505 510

Leu Ala Thr Gly Arg Ala Ala Met Glu His Arg Ala Val Val Val Ala  
515 520 525

Asp Asp Arg Asp Ser Phe Val Ala Gly Leu Thr Ala Leu Ala Ala Gly  
530 535 540

Val Pro Ala Ala Asn Val Val Gln Gly Ala Ala Asp Cys Lys Gly Lys  
545 550 555 560

Val Ala Phe Val Phe Pro Gly Gln Gly Ser His Trp Gln Gly Met Ala  
565 570 575

Arg Glu Leu Ser Glu Ser Ser Pro Val Phe Arg Arg Lys Leu Ala Glu  
 580 585 590  
 Cys Ala Ala Ala Thr Ala Pro Tyr Val Asp Trp Ser Leu Leu Gly Val  
 595 600 605  
 Leu Arg Gly Asp Pro Asp Ala Pro Ala Leu Asp Arg Asp Asp Val Ile  
 610 615 620  
 Gln Leu Ala Leu Phe Ala Met Met Val Ser Leu Ala Glu Leu Trp Arg  
 625 630 635 640  
 Ser Cys Gly Val Glu Pro Ala Ala Val Val Gly His Ser Gln Gly Glu  
 645 650 655  
 Ile Ala Ala Ala His Val Ala Gly Ala Leu Ser Leu Thr Asp Ala Val  
 660 665 670  
 Arg Ile Ile Ala Ala Arg Cys Asp Ala Val Ser Ala Leu Thr Gly Lys  
 675 680 685  
 Gly Gly Met Leu Ala Ile Ala Leu Pro Glu Ser Ala Val Val Lys Arg  
 690 695 700  
 Ile Ala Gly Leu Pro Glu Leu Thr Val Ala Ala Val Asn Gly Pro Gly  
 705 710 715 720  
 Ser Thr Val Val Ser Gly Glu Pro Ser Ala Leu Glu Arg Leu Gln Thr  
 725 730 735  
 Glu Leu Thr Ala Glu Asn Val Gln Thr Arg Arg Val Gly Ile Asp Tyr  
 740 745 750  
 Ala Ser His Ser Pro Gln Ile Ala Gln Val Gln Gly Arg Leu Leu Asp  
 755 760 765  
 Arg Leu Gly Glu Val Gly Ser Glu Pro Ala Glu Ile Ala Phe Tyr Ser  
 770 775 780  
 Thr Val Thr Gly Glu Arg Thr Asp Thr Gly Arg Leu Asp Ala Asp Tyr  
 785 790 795 800  
 Trp Tyr Gln Asn Leu Arg Gln Pro Val Arg Phe Gln Gln Thr Val Ala  
 805 810 815  
 Arg Met Ala Asp Gln Gly Tyr Arg Phe Phe Val Glu Val Ser Pro His  
 820 825 830

Pro Leu Leu Thr Ala Gly Ile Gln Glu Thr Leu Glu Ala Ala Asp Ala  
 835 340 845

Gly Gly Val Val Val Gly Ser Leu Arg Arg Gly Glu Gly Gly Ser Arg  
 850 855 860

Arg Trp Leu Thr Ser Leu Ala Glu Cys Gln Val Arg Gly Leu Pro Val  
 865 870 875 880

Asn Trp Glu Gln Val Phe Leu Asn Thr Gly Ala Arg Arg Val Pro Leu  
 885 890 895

Pro Thr Tyr Pro Phe Gln Arg Gln Arg Tyr Trp Leu Glu Ser Ala Glu  
 900 905 910

Tyr Asp Ala Gly Asp Leu Gly Ser Val Gly Leu Leu Ser Ala Glu His  
 915 920 925

Pro Leu Leu Gly Ala Ala Val Thr Leu Ala Asp Ala Gly Gly Phe Leu  
 930 935 940

Leu Thr Gly Lys Leu Ser Val Lys Thr Gln Pro Trp Leu Ala Asp His  
 945 950 955 960

Val Val Gly Gly Ala Ile Leu Leu Pro Gly Thr Ala Phe Val Glu Met  
 965 970 975

Leu Ile Arg Ala Ala Asp Gln Val Gly Cys Asp Leu Ile Glu Glu Leu  
 980 985 990

Ser Leu Thr Thr Pro Leu Val Leu Pro Ala Thr Gly Ala Val Gln Val  
 995 1000 1005

Gln Ile Ala Val Gly Gly Pro Asp Glu Ala Gly Arg Arg Ser Val Arg  
 1010 1015 1020

Val His Ser Cys Arg Asp Asp Ala Val Pro Gln Asp Ser Trp Thr Cys  
 1025 1030 1035 1040

His Ala Thr Gly Thr Leu Thr Ser Ser Asp His Gln Asp Ala Gly Gln  
 1045 1050 1055

Gly Pro Asp Gly Ile Trp Pro Pro Asn Asp Ala Val Ala Val Pro Leu  
 1060 1065 1070

Asp Ser Phe Tyr Ala Arg Ala Ala Glu Arg Gly Phe Asp Phe Gly Pro  
 1075 1080 1085

Ala Phe Gln Gly Leu Gln Ala Ala Trp Lys Arg Gly Asp Glu Ile Phe  
1090 1095 1100

Ala Glu Val Gly Leu Pro Thr Ala His Arg Glu Asp Ala Gly Arg Phe  
105 1110 1115 1120

Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu Gln Ala Leu Gly Ala  
1125 1130 1135

Ala Glu Glu Asp Pro Asp Glu Gly Trp Leu Pro Phe Ala Trp Gln Gly  
1140 1145 1150

Val Ser Leu Lys Ala Thr Gly Ala Leu Ser Leu Arg Val His Leu Val  
1155 1160 1165

Pro Ala Gly Ala Asn Ala Val Ser Val Phe Thr Thr Asp Thr Thr Gly  
1170 1175 1180

Gln Ala Val Leu Ser Ile Asp Ser Leu Val Leu Arg Gln Ile Ser Asp  
185 1190 1195 1200

Lys Gln Leu Ala Ala Ala Arg Ala Met Glu His Glu Ser Leu Phe Arg  
1205 1210 1215

Val Asp Trp Lys Arg Ile Ser Pro Gly Ala Ala Lys Pro Val Ser Trp  
1220 1225 1230

Ala Val Ile Gly Asn Asp Glu Leu Ala Arg Ala Cys Gly Ser Ala Leu  
1235 1240 1245

Gly Thr Glu Leu His Pro Asp Leu Thr Gly Leu Ala Asp Pro Pro Pro  
1250 1255 1260

Asp Val Val Val Val Pro Cys Gly Ala Ser Arg Gln Asp Leu Asp Val  
265 1270 1275 1280

Ala Ser Glu Ala Arg Ala Ala Thr Gln Arg Met Leu Asp Leu Ile Gln  
1285 1290 1295

Asp Trp Leu Ala Ala Ala Arg Phe Ala Gly Ser Arg Leu Val Val Val  
1300 1305 1310

Thr Cys Gly Ala Ala Ser Thr Gly Pro Ala Glu Gly Val Ser Asp Leu  
1315 1320 1325

Val His Ala Ala Ser Trp Gly Leu Leu Arg Ser Ala Gln Ser Glu Asn  
1330 1335 1340

Pro Asp Arg Phe Val Leu Val Asp Val Asp Gly Thr Ala Glu Ser Trp  
 345 1350 1355 1360

Arg Ala Leu Ala Ala Ala Val Arg Ser Gly Glu Pro Gln Leu Ala Leu  
 1365 1370 1375

Arg Ala Gly Glu Val Arg Val Pro Arg Leu Ala Arg Cys Val Ala Ala  
 1380 1385 1390

Glu Asp Ser Arg Ile Pro Val Pro Gly Ala Asp Gly Thr Val Leu Ile  
 1395 1400 1405

Ser Gly Gly Thr Gly Leu Leu Gly Gly Leu Val Ala Arg His Leu Val  
 1410 1415 1420

Ala Glu Arg Gly Val Arg Arg Leu Val Leu Ala Gly Arg Arg Gly Trp  
 425 1430 1435 1440

Ser Ala Pro Gly Val Thr Asp Leu Val Asp Glu Leu Val Gly Leu Gly  
 1445 1450 1455

Ala Ala Val Glu Val Ala Ser Cys Asp Val Gly Asp Arg Ala Gln Leu  
 1460 1465 1470

Asp Arg Leu Leu Thr Thr Ile Ser Ala Glu Phe Pro Leu Arg Gly Val  
 1475 1480 1485

Val His Ala Ala Gly Ala Leu Ala Asp Gly Val Val Glu Ser Leu Thr  
 1490 1495 1500

Pro Glu His Val Ala Lys Val Phe Gly Pro Lys Ala Ala Gly Ala Trp  
 505 1510 1515 1520

His Leu His Glu Leu Thr Leu Asp Leu Asp Leu Ser Phe Phe Val Leu  
 1525 1530 1535

Phe Ser Ser Phe Ser Gly Val Ala Gly Ala Ala Gly Gln Gly Asn Tyr  
 1540 1545 1550

Ala Ala Ala Asn Ala Phe Leu Asp Gly Leu Ala Gln His Arg Arg Thr  
 1555 1560 1565

Ala Gly Leu Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro  
 1570 1575 1580

Ser Gly Met Thr Gly Ala Leu Asp Ala Ala Gly Arg Ser Arg Ile Ala  
 585 1590 1595 1600

Arg Thr Asn Pro Pro Met Ser Ala Pro Asp Gly Leu Arg Leu Phe Glu  
 1605 1610 1615  
 Met Ala Phe Arg Val Pro Gly Glu Ser Leu Leu Val Pro Val His Val  
 1620 1625 1630  
 Asp Leu Asn Ala Leu Arg Ala Asp Ala Ala Asp Gly Gly Val Pro Ala  
 1635 1640 1645  
 Leu Leu Arg Asp Leu Val Pro Ala Pro Val Arg Arg Ser Ala Val Asn  
 1650 1655 1660  
 Glu Ser Ala Asp Val Asn Gly Leu Val Gly Arg Leu Arg Arg Leu Pro  
 665 1670 1675 1680  
 Asp Leu Asp Gln Glu Thr Gln Leu Leu Gly Leu Val Arg Glu His Val  
 1685 1690 1695  
 Ser Ala Val Leu Gly His Ser Gly Ala Val Glu Val Gly Ala Asp Arg  
 1700 1705 1710  
 Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Phe Arg  
 1715 1720 1725  
 Asn Arg Leu Gly Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val  
 1730 1735 1740  
 Phe Asp Tyr Pro Thr Pro Arg Ala Leu Val Arg Phe Leu Leu Asp Lys  
 745 1750 1755 1760  
 Leu Ile Gly Gly Val Glu Ala Pro Thr Pro Ala Pro Ala Ala Val Ala  
 1765 1770 1775  
 Ala Val Thr Ala Asp Asp Pro Val Val Ile Val Gly Met Gly Cys Arg  
 1780 1785 1790  
 Tyr Pro Gly Gly Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala  
 1795 1800 1805  
 Gly Gly Leu Asp Ala Val Ala Glu Phe Pro Asp Asp Arg Gly Trp Asp  
 1810 1815 1820  
 Gln Ala Gly Leu Phe Asp Pro Asp Pro Asp Arg Leu Gly Thr Ser Tyr  
 825 1830 1835 1840  
 Val Cys Glu Gly Gly Phe Leu Arg Asp Ala Ala Glu Phe Asp Ala Gly  
 1845 1850 1855

Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln  
1860 1865 1870  
Arg Leu Leu Leu Glu Val Ala Trp Glu Thr Val Glu Arg Ala Gly Ile  
1875 1880 1885  
Asp Pro Leu Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu  
1890 1895 1900  
Met His His Asp Tyr Gly Ala Arg Phe Ile Thr Arg Ala Pro Glu Gly  
905 1910 1915 1920  
Phe Glu Gly Tyr Leu Gly Asn Gly Ser Ala Gly Gly Val Phe Ser Gly  
1925 1930 1935  
Arg Val Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp  
1940 1945 1950  
Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ala  
1955 1960 1965  
Leu Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val  
1970 1975 1980  
Met Ala Thr Pro Gly Met Phe Val Glu Phe Ser Arg Gln Arg Gly Leu  
985 1990 1995 2000  
Ala Ala Asp Gly Arg Cys Lys Ser Phe Ala Ala Ala Ala Asp Gly Thr  
2005 2010 2015  
Gly Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp  
2020 2025 2030  
Ala Arg Arg Asn Gly His Ala Val Leu Ala Val Val Arg Gly Ser Ala  
2035 2040 2045  
Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro  
2050 2055 2060  
Ser Gln Gln Arg Val Ile Thr Gln Ala Leu Ala Ser Ala Gly Leu Ser  
065 2070 2075 2080  
Val Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu  
2085 2090 2095  
Gly Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Gln Gly  
2100 2105 2110

Arg Asp Ser Asp Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile  
 2115 2120 2125

Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val  
 2130 2135 2140

Met Ala Met Arg His Gly Gln Leu Pro Ala Thr Leu His Val Asp Glu  
 145 2150 2155 2160

Pro Thr Ser Glu Val Asp Trp Ser Ala Gly Asp Val Gln Leu Leu Thr  
 2165 2170 2175

Glu Asn Thr Pro Trp Pro Gly Asn Ser His Pro Arg Arg Val Gly Val  
 2180 2185 2190

Ser Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln  
 2195 2200 2205

Ala Ser Lys Thr Pro Asp Glu Thr Ala Asp Lys Ser Gly Pro Asp Ser  
 2210 2215 2220

Glu Ser Thr Val Asp Leu Pro Ala Val Pro Leu Ile Val Ser Gly Arg  
 225 2230 2235 2240

Thr Pro Ala Ala Leu Ser Ala Gln Ala Ser Ala Leu Leu Ser Tyr Leu  
 2245 2250 2255

Gly Glu Arg Gly Asp Ile Ser Thr Leu Asp Ala Ala Phe Ser Leu Ala  
 2260 2265 2270

Ser Ser Arg Ala Ala Leu Glu Glu Arg Ala Val Val Leu Gly Ala Asp  
 2275 2280 2285

Arg Glu Thr Leu Leu Ser Gly Leu Glu Ala Leu Ala Ser Gly Arg Glu  
 2290 2295 2300

Ala Ser Gly Val Val Ser Gly Ser Pro Val Ser Gly Gly Val Gly Phe  
 305 2310 2315 2320

Val Phe Ala Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Gly Leu  
 2325 2330 2335

Tyr Ser Val Phe Pro Val Phe Ala Asp Ala Phe Asp Glu Ala Cys Ala  
 2340 2345 2350

Gly Leu Asp Ala His Leu Gly Gln Asp Val Gly Val Arg Asp Val Val  
 2355 2360 2365



Phe Gly Ser Asp Gly Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser  
2370 2375 2380

Gly Leu Phe Ala Leu Gln Val Gly Leu Leu Ser Leu Leu Gly Ser Trp  
385 2390 2395 2400

Gly Val Arg Pro Gly Val Val Leu Gly His Ser Val Gly Glu Phe Ala  
2405 2410 2415

Ala Ala Val Ala Ala Gly Val Leu Ser Leu Pro Asp Ala Ala Arg Met  
2420 2425 2430

Val Ala Gly Arg Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala  
2435 2440 2445

Met Leu Ala Val Ala Ala Gly Glu Glu Gln Leu Arg Pro Leu Leu Ala  
2450 2455 2460

Asp Arg Val Asp Gly Ala Gly Ile Ala Ala Val Asn Ala Pro Glu Ser  
465 2470 2475 2480

Val Val Leu Ser Gly Asp Arg Glu Val Leu Asp Asp Ile Ala Gly Ala  
2485 2490 2495

Leu Asp Gly Gln Gly Ile Arg Trp Arg Arg Leu Arg Val Ser His Ala  
2500 2505 2510

Phe His Ser Tyr Arg Met Asp Pro Met Leu Gln Glu Phe Ala Glu Ile  
2515 2520 2525

Ala Arg Ser Val Asp Tyr Arg Arg Gly Asp Leu Pro Val Val Ser Thr  
2530 2535 2540

Leu Thr Gly Glu Leu Asp Thr Ala Gly Val Met Ala Thr Pro Glu Tyr  
545 2550 2555 2560

Trp Val Arg Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Arg  
2565 2570 2575

Val Leu Ala Gln Gln Gly Val Ala Thr Ile Phe Glu Leu Gly Pro Asp  
2580 2585 2590

Ala Thr Leu Ser Ala Leu Ile Pro Asp Cys His Ser Trp Ala Asp Gln  
2595 2600 2605

Ala Met Pro Ile Pro Met Leu Arg Lys Asp Arg Thr Glu Thr Glu Thr  
2610 2615 2620



Gly Asp Glu Val Leu Ala Glu Ile Ala Pro Pro Ala Glu Ala Asp Ala  
2885 2890 2895

Met Ala Ser Gly Tyr Leu Leu Asp Pro Ala Leu Leu Asp Ala Ala Leu  
2900 2905 2910

Gln Ala Ser Ala Leu Gly Asp Arg Pro Glu Gln Gly Gly Ala Trp Leu  
2915 2920 2925

Pro Phe Ser Phe Thr Gly Val Glu Leu Ser Ala Pro Ala Gly Thr Ile  
2930 2935 2940

Ser Arg Val Arg Leu Glu Thr Arg Arg Pro Asp Ala Ile Ser Val Ala  
2945 2950 2955 2960

Val Met Asp Glu Ser Gly Arg Leu Leu Ala Ser Ile Asp Ser Leu Arg  
2965 2970 2975

Leu Arg Ser Val Ser Ser Gly Gln Leu Ala Asn Arg Asp Ala Val Arg  
2980 2985 2990

Asp Ala Leu Phe Glu Val Thr Trp Glu Pro Val Ala Thr Gln Ser Thr  
2995 3000 3005

Glu Pro Gly Arg Trp Ala Leu Leu Gly Asp Thr Ala Cys Gly Lys Asp  
3010 3015 3020

Asp Leu Ile Lys Leu Ala Thr Asp Ser Ala Asp Arg Cys Ala Asp Leu  
3025 3030 3035 3040

Ala Ala Leu Ala Glu Lys Leu Asp Ser Ser Ala Leu Val Pro Asp Val  
3045 3050 3055

Val Val Tyr Cys Ala Gly Glu Gln Ala Asp Pro Gly Thr Gly Ala Ala  
3060 3065 3070

Ala Leu Ala Glu Thr Gln Gln Thr Leu Ala Leu Leu Gln Ala Trp Leu  
3075 3080 3085

Ala Glu Pro Arg Leu Ala Glu Ala Arg Leu Val Val Val Thr Cys Ala  
3090 3095 3100

Ala Val Thr Thr Ala Pro Ser Asp Gly Ala Ser Glu Leu Ala His Ala  
3105 3110 3115 3120

Pro Leu Trp Gly Leu Leu Arg Ala Ala Gln Val Glu Asn Pro Gly Gln  
3125 3130 3135

Phe Val Leu Ala Asp Val Asp Gly Thr Ala Glu Ser Trp Arg Ala Leu  
 3140 3145 3150  
 Pro Ser Ala Leu Gly Ser Met Glu Pro Gln Leu Ala Leu Arg Lys Gly  
 3155 3160 3165  
 Ala Val Arg Ala Pro Arg Leu Ala Ser Val Ala Gly Gln Ile Asp Val  
 3170 3175 3180  
 Pro Ala Val Val Ala Asp Pro Asp Arg Thr Val Leu Ile Ser Gly Gly  
 185 3190 3195 3200  
 Thr Gly Leu Leu Gly Gly Ala Val Ala Arg His Leu Val Thr Glu Arg  
 3205 3210 3215  
 Gly Val Arg Arg Leu Val Leu Thr Gly Arg Arg Gly Trp Asp Ala Pro  
 3220 3225 3230  
 Gly Ile Thr Glu Leu Val Gly Glu Leu Asn Gly Leu Gly Ala Val Val  
 3235 3240 3245  
 Asp Val Val Ala Cys Asp Val Ala Asp Arg Ala Asp Leu Glu Ser Leu  
 3250 3255 3260  
 Leu Ala Ala Val Pro Ala Glu Phe Pro Leu Cys Gly Val Val His Ala  
 265 3270 3275 3280  
 Ala Gly Ala Leu Ala Asp Gly Val Ile Glu Ser Leu Ser Pro Asp Asp  
 3285 3290 3295  
 Val Gly Ala Val Phe Gly Pro Lys Ala Ala Gly Ala Trp Asn Leu His  
 3300 3305 3310  
 Glu Leu Thr Arg Asp Thr Asp Leu Ser Phe Phe Ala Leu Phe Ser Ser  
 3315 3320 3325  
 Leu Ser Gly Val Ala Gly Ala Pro Gly Gln Gly Asn Tyr Ala Ala Ala  
 3330 3335 3340  
 Asn Ala Phe Leu Asp Ala Leu Ala His Tyr Arg Arg Ser Gln Gly Leu  
 345 3350 3355 3360  
 Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Gln Pro Ser Gly Met  
 3365 3370 3375  
 Thr Glu Thr Leu Ser Glu Val Asp Arg Ser Arg Ile Ala Arg Ala Asn  
 3380 3385 3390

Pro Pro Leu Ser Thr Lys Glu Gly Leu Arg Leu Phe Asp Ala Gly Leu  
3395 3400 3405

Ala Leu Asp Arg Ala Ala Val Val Pro Ala Lys Leu Asp Arg Thr Phe  
3410 3415 3420

Leu Ala Glu Gln Ala Arg Ser Gly Ser Leu Pro Ala Leu Leu Thr Ala  
425 3430 3435 3440

Leu Val Pro Pro Ile Arg Arg Asn Arg Arg Ala Ser Gly Thr Glu Leu  
3445 3450 3455

Ala Asp Glu Gly Thr Leu Leu Gly Val Val Arg Glu His Ala Ala Ala  
3460 3465 3470

Val Leu Gly Tyr Ser Ser Ala Ala Asp Val Gly Val Glu Arg Ala Phe  
3475 3480 3485

Arg Asp Leu Gly Phe Asp Ser Leu Ser Gly Val Glu Leu Arg Asn Arg  
3490 3495 3500

Leu Ala Gly Val Leu Gly Val Arg Leu Pro Ala Thr Ala Val Phe Asp  
505 3510 3515 3520

Tyr Pro Thr Pro Arg Ala Leu Ala Arg Phe Leu His Gln Glu Leu Ala  
3525 3530 3535

Asp Glu Ile Ala Thr Thr Pro Ala Pro Val Thr Thr Thr Arg Ala Pro  
3540 3545 3550

Val Ala Glu Asp Asp Leu Val Ala Ile Val Gly Met Gly Cys Arg Phe  
3555 3560 3565

Pro Gly Gln Val Ser Ser Pro Glu Glu Leu Trp Arg Leu Val Ala Gly  
3570 3575 3580

Gly Val Asp Ala Val Ala Asp Phe Pro Ala Asp Arg Gly Trp Asp Leu  
585 3590 3595 3600

Ala Gly Leu Phe Asp Pro Asp Pro Glu Arg Ala Gly Lys Thr Tyr Val  
3605 3610 3615

Arg Glu Gly Ala Phe Leu Thr Asp Ala Asp Arg Phe Asp Ala Gly Phe  
3620 3625 3630

Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg  
3635 3640 3645

Leu Leu Leu Glu Leu Ser Trp Glu Ala Ile Glu Arg Ala Gly Ile Asp  
 3650 3655 3660  
 Pro Gly Ser Leu Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Leu Met  
 665 3670 3675 3680  
 Tyr His Asp Tyr Gly Ala Arg Phe Ala Ser Arg Ala Pro Glu Gly Phe  
 3685 3690 3695  
 Glu Gly Tyr Leu Gly Asn Gly Ser Ala Gly Ser Val Ala Ser Gly Arg  
 3700 3705 3710  
 Ile Ala Tyr Ser Phe Gly Phe Glu Gly Pro Ala Val Thr Val Asp Thr  
 3715 3720 3725  
 Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Gly Gln Ser Leu  
 3730 3735 3740  
 Arg Ser Gly Glu Cys Asp Leu Ala Leu Ala Gly Gly Val Thr Val Met  
 745 3750 3755 3760  
 Ser Thr Pro Gly Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala  
 3765 3770 3775  
 Pro Asp Gly Arg Cys Lys Ser Phe Ala Glu Ser Ala Asp Gly Thr Gly  
 3780 3785 3790  
 Trp Gly Glu Gly Ala Gly Leu Val Leu Leu Glu Arg Leu Ser Asp Ala  
 3795 3800 3805  
 Arg Arg Asn Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val  
 3810 3815 3820  
 Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser  
 825 3830 3835 3840  
 Gln Gln Arg Val Ile Gln Gln Ala Leu Ala Ser Ala Gly Leu Ser Val  
 3845 3850 3855  
 Ser Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Arg Leu Gly  
 3860 3865 3870  
 Asp Pro Ile Glu Ala Gln Ala Leu Ile Ala Thr Tyr Gly Arg Asp Arg  
 3875 3880 3885  
 Asp Pro Gly Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly  
 3890 3895 3900

His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met  
905 3910 3915 3920

Ala Met Arg His Gly Gln Leu Pro Arg Thr Leu His Val Asp Ala Pro  
3925 3930 3935

Ser Ser Gln Val Asp Trp Ser Ala Gly Arg Val Gln Leu Leu Thr Glu  
3940 3945 3950

Asn Thr Pro Trp Pro Asp Ser Gly Arg Pro Cys Arg Val Gly Val Ser  
3955 3960 3965

Ser Phe Gly Ile Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ser  
3970 3975 3980

Thr Gly Gln Met Asp Gln Ala Ala Glu Pro Asp Ser Ser Pro Val Leu  
985 3990 3995 4000

Asp Val Pro Val Val Pro Trp Val Val Ser Gly Lys Thr Pro Glu Ala  
4005 4010 4015

Leu Ser Ala Gln Ala Ala Thr Leu Ala Thr Tyr Leu Asp Gln Asn Val  
4020 4025 4030

Asp Val Ser Pro Leu Asp Val Gly Ile Ser Leu Ala Val Thr Arg Ser  
4035 4040 4045

Ala Leu Asp Glu Arg Ala Val Val Leu Gly Ser Asp Arg Asp Thr Leu  
4050 4055 4060

Leu Ser Gly Leu Asn Ala Leu Ala Ala Gly His Glu Ala Ala Gly Val  
065 4070 4075 4080

Val Thr Gly Pro Val Gly Ile Gly Gly Arg Thr Gly Phe Val Phe Ala  
4085 4090 4095

Gly Gln Gly Gly Gln Trp Leu Gly Met Gly Arg Arg Leu Tyr Ser Glu  
4100 4105 4110

Phe Pro Ala Phe Ala Gly Ala Phe Asp Glu Ala Cys Ala Glu Leu Asp  
4115 4120 4125

Ala Asn Leu Gly Arg Glu Val Gly Val Arg Asp Val Val Phe Gly Ser  
4130 4135 4140

Asp Glu Ser Leu Leu Asp Arg Thr Leu Trp Ala Gln Ser Gly Leu Phe  
145 4150 4155 4160

Ala Leu Gln Val Gly Leu Trp Glu Leu Leu Gly Thr Trp Gly Val Arg  
4165 4170 4175

Pro Ser Val Val Leu Gly His Ser Val Gly Glu Leu Ala Ala Ala Phe  
4180 4185 4190

Ala Ala Gly Val Leu Ser Met Ala Glu Ala Ala Arg Leu Val Ala Gly  
4195 4200 4205

Arg Ala Arg Leu Met Gln Ala Leu Pro Ser Gly Gly Ala Met Leu Ala  
4210 4215 4220

Val Ser Ala Thr Glu Ala Arg Val Gly Pro Leu Leu Asp Gly Val Arg  
225 4230 4235 4240

Asp Arg Val Gly Val Ala Ala Val Asn Ala Pro Gly Ser Val Val Leu  
4245 4250 4255

Ser Gly Asp Arg Asp Val Leu Asp Gly Ile Ala Gly Arg Leu Asp Gly  
4260 4265 4270

Gln Gly Ile Arg Ser Arg Trp Leu Arg Val Ser His Ala Phe His Ser  
4275 4280 4285

His Arg Met Asp Pro Met Leu Ala Glu Phe Ala Glu Leu Ala Arg Ser  
4290 4295 4300

Val Asp Tyr Arg Ser Pro Arg Leu Pro Ile Val Ser Thr Leu Thr Gly  
305 4310 4315 4320

Asn Leu Asp Asp Val Gly Val Met Ala Thr Pro Glu Tyr Trp Val Arg  
4325 4330 4335

Gln Val Arg Glu Pro Val Arg Phe Ala Asp Gly Val Gln Ala Leu Val  
4340 4345 4350

Asp Gln Gly Val Asp Thr Ile Val Glu Leu Gly Pro Asp Gly Ala Leu  
4355 4360 4365

Ser Ser Leu Val Gln Glu Cys Val Ala Glu Ser Gly Arg Ala Thr Gly  
4370 4375 4380

Ile Pro Leu Val Arg Arg Asp Arg Asp Glu Val Arg Thr Val Leu Asp  
385 4390 4395 4400

Ala Leu Ala Gln Thr His Thr Arg Gly Gly Ala Val Asp Trp Gly Ser  
4405 4410 4415



Phe Phe Ala Gly Thr Arg Ala Thr Gln Val Asp Leu Pro Thr Tyr Ala  
4420 4425 4430

Phe Gln Arg Gln Arg Tyr Trp Leu Glu Pro Ser Asp Ser Gly Asp Val  
4435 4440 4445

Thr Gly Val Gly Leu Thr Gly Ala Glu His Pro Leu Leu Gly Ala Val  
4450 4455 4460

Val Pro Val Ala Gly Gly Asp Glu Val Leu Leu Thr Gly Arg Leu Ser  
4465 4470 4475 4480

Val Gly Thr His Pro Trp Leu Ala Glu His Arg Val Leu Gly Glu Val  
4485 4490 4495

Val Val Pro Gly Thr Ala Leu Leu Glu Met Ala Trp Arg Ala Gly Ser  
4500 4505 4510

Gln Val Gly Cys Glu Arg Val Glu Glu Leu Thr Leu Glu Ala Pro Leu  
4515 4520 4525

Val Leu Pro Glu Arg Gly Ala Ala Ala Val Gln Leu Ala Val Gly Ala  
4530 4535 4540

Pro Asp Glu Ala Gly Arg Arg Ser Leu Gln Leu Tyr Ser Arg Gly Ala  
4545 4550 4555 4560

Asp Glu Asp Gly Asp Trp Arg Arg Ile Ala Ser Gly Leu Leu Ala Gln  
4565 4570 4575

Ala Asn Ala Val Pro Pro Ala Asp Ser Thr Ala Trp Pro Pro Asp Gly  
4580 4585 4590

Ala Gly Gln Val Asp Leu Ala Glu Phe Tyr Glu Arg Leu Ala Glu Arg  
4595 4600 4605

Gly Leu Thr Tyr Gly Pro Val Phe Gln Gly Leu Arg Ala Ala Trp Arg  
4610 4615 4620

His Gly Asp Asp Ile Phe Ala Glu Leu Ala Gly Ser Pro Asp Ala Ser  
4625 4630 4635 4640

Gly Phe Gly Ile His Pro Ala Leu Leu Asp Ala Ala Leu His Ala Met  
4645 4650 4655

Ala Leu Gly Ala Ser Pro Asp Ser Glu Ala Arg Leu Pro Phe Ser Trp  
4660 4665 4670

Arg Gly Ala Gln Leu Tyr Arg Ala Glu Gly Ala Ala Leu Arg Val Arg  
 4675 4680 4685

Leu Ser Pro Leu Gly Ser Gly Ala Val Ser Leu Thr Leu Val Asp Ala  
 4690 4695 4700

Thr Gly Arg Arg Val Ala Ala Val Glu Ser Leu Ser Thr Arg Pro Val  
 705 4710 4715 4720

Ser Thr Asp Gln Ile Gly Ala Gly Arg Gly Asp Gln Glu Arg Leu Leu  
 4725 4730 4735

His Val Glu Trp Val Arg Ser Ala Glu Ser Ala Gly Met Ser Leu Thr  
 4740 4745 4750

Ser Cys Ala Val Val Gly Leu Gly Glu Pro Glu Trp His Ala Ala Leu  
 4755 4760 4765

Lys Thr Thr Gly Val Gln Val Glu Ser His Ala Asp Leu Ala Ser Leu  
 4770 4775 4780

Ala Thr Glu Val Ala Lys Arg Gly Ser Ala Pro Gly Ala Val Ile Val  
 785 4790 4795 4800

Pro Cys Pro Arg Pro Arg Ala Met Gln Glu Leu Pro Thr Ala Ala Arg  
 4805 4810 4815

Arg Ala Thr Gln Gln Ala Met Ala Met Leu Gln Gln Trp Leu Ala Asp  
 4820 4825 4830

Asp Arg Phe Val Ser Thr Arg Leu Ile Leu Leu Thr His Arg Ala Val  
 4835 4840 4845

Ser Ala Val Ala Gly Glu Asp Val Leu Asp Leu Val His Ala Pro Leu  
 4850 4855 4860

Trp Gly Leu Val Arg Ser Ala Gln Ala Glu His Pro Asp Arg Phe Ala  
 865 4870 4875 4880

Leu Ile Asp Met Asp Asp Glu Arg Ala Ser Gln Thr Ala Leu Ala Glu  
 4885 4890 4895

Ala Leu Thr Ala Gly Glu Ala Gln Leu Ala Val Arg Ser Gly Val Val  
 4900 4905 4910

Leu Ala Pro Arg Leu Gly Gln Val Lys Val Ser Gly Gly Glu Ala Phe  
 4915 4920 4925

Arg Trp Asp Glu Gly Thr Val Leu Val Thr Gly Gly Thr Gly Gly Leu  
 4930 4935 4940

Gly Ala Leu Leu Ala Arg His Leu Val Ser Ala His Gly Val Arg His  
 945 4950 4955 4960

Leu Leu Leu Ala Ser Arg Arg Gly Leu Ala Ala Pro Gly Ala Asp Glu  
 4965 4970 4975

Leu Val Ala Glu Leu Glu Gln Ala Gly Ala Asp Val Ala Val Val Ala  
 4980 4985 4990

Cys Asp Ser Ala Asp Arg Asp Ser Leu Ala Arg Leu Val Ala Ser Val  
 4995 5000 5005

Pro Ala Glu Asn Pro Leu Arg Val Val Val His Ala Ala Gly Val Leu  
 5010 5015 5020

Asp Asp Gly Val Leu Met Ser Met Ser Pro Glu Arg Leu Asp Ala Val  
 5025 5030 5035 5040

Leu Arg Pro Lys Val Asp Ala Ala Trp Tyr Leu His Glu Leu Thr Arg  
 5045 5050 5055

Glu Leu Gly Leu Ser Ala Phe Val Leu Phe Ser Ser Val Ala Gly Leu  
 5060 5065 5070

Phe Gly Gly Ala Gly Gln Ser Asn Tyr Ala Ala Gly Asn Ala Phe Leu  
 5075 5080 5085

Asp Ala Leu Ala His Cys Arg Gln Ala Gln Gly Leu Pro Ala Leu Ser  
 5090 5095 5100

Leu Ala Ser Gly Leu Trp Ala Ser Ile Asp Gly Met Ala Gly Asp Leu  
 5105 5110 5115 5120

Ala Ala Ala Asp Val Glu Arg Leu Ser Arg Ala Gly Ile Gly Pro Leu  
 5125 5130 5135

Ser Ala Pro Gly Gly Leu Ala Leu Phe Asp Ala Ala Val Gly Ser Asp  
 5140 5145 5150

Glu Pro Leu Leu Ala Pro Val Arg Leu Asp Val Glu Ala Leu Arg Val  
 5155 5160 5165

Gln Ala Arg Ser Val Gln Thr Arg Ile Pro Glu Met Leu His Gly Met  
 5170 5175 5180

Ala Met Gly Pro Ser Arg Arg Thr Pro Phe Thr Ser Arg Val Glu Pro  
185 5190 5195 5200

Leu His Glu Arg Leu Ala Gly Leu Ser Glu Gly Glu Arg Arg Gln Gln  
5205 5210 5215

Val Leu Gln Arg Val Arg Ala Asp Ile Ala Val Val Leu Gly His Gly  
5220 5225 5230

Arg Ser Ser Asp Val Asp Ile Glu Lys Pro Leu Ala Glu Leu Gly Phe  
5235 5240 5245

Asp Ser Leu Thr Ala Ile Glu Leu Arg Asn Arg Leu Ala Thr Ala Thr  
5250 5255 5260

Gly Leu Arg Leu Pro Ala Thr Leu Ala Phe Asp His Gly Thr Ala Ala  
265 5270 5275 5280

Ala Leu Ala Gln His Val Cys Ala Gln Leu Gly Thr Ala Thr Ala Pro  
5285 5290 5295

Ala Pro Arg Arg Thr Asp Asp Asn Asp Ala Thr Glu Pro Val Arg Ser  
5300 5305 5310

Leu Phe Gln Gln Ala Tyr Ala Ala Gly Arg Ile Leu Asp Gly Met Asp  
5315 5320 5325

Leu Val Lys Val Ala Ala Gln Leu Arg Pro Val Phe Gly Ser Pro Gly  
5330 5335 5340

Glu Leu Glu Ser Leu Pro Lys Pro Val Gln Leu Ser Arg Gly Pro Glu  
345 5350 5355 5360

Glu Leu Ala Leu Val Cys Met Pro Ala Leu Ile Gly Met Pro Pro Ala  
5365 5370 5375

Gln Gln Tyr Ala Arg Ile Ala Ala Gly Phe Arg Asp Val Arg Asp Val  
5380 5385 5390

Ser Val Ile Pro Met Pro Gly Phe Ile Ala Gly Glu Pro Leu Pro Ser  
5395 5400 5405

Ala Ile Glu Val Ala Val Arg Thr Gln Ala Glu Ala Val Leu Gln Glu  
5410 5415 5420

Phe Ala Gly Gly Ser Phe Val Leu Val Gly His Ser Ser Gly Gly Trp  
425 5430 5435 5440

Leu Ala His Glu Val Ala Gly Glu Leu Glu Arg Arg Gly Val Val Pro  
 5445 5450 5455

Ala Gly Val Val Leu Leu Asp Thr Tyr Ile Pro Gly Glu Ile Thr Pro  
 5460 5465 5470

Arg Phe Ser Val Ala Met Ala His Arg Thr Tyr Glu Lys Leu Ala Thr  
 5475 5480 5485

Phe Thr Asp Met Gln Asp Val Gly Ile Thr Ala Met Gly Gly Tyr Phe  
 5490 5495 5500

Arg Met Phe Thr Glu Trp Thr Pro Thr Pro Ile Gly Ala Pro Thr Leu  
 505 5510 5515 5520

Phe Val Arg Thr Glu Asp Cys Val Ala Asp Pro Glu Gly Arg Pro Trp  
 5525 5530 5535

Thr Asp Asp Ser Trp Arg Pro Gly Trp Thr Leu Ala Asp Ala Thr Val  
 5540 5545 5550

Gln Val Pro Gly Asp His Phe Ser Met Met Asp Glu His Ala Gly Ser  
 5555 5560 5565

Thr Ala Gln Ala Val Ala Ser Trp Leu Asp Lys Leu Asn Gln Arg Thr  
 5570 5575 5580

Ala Arg Gln Arg  
 585

<210> 51

<211> 2363

<212> DNA

<213> *Saccharopolyspora spinosa*

<400> 51

cccgggcccgg ccgtgggccc cgggtccgcg gggcggaacg ggtctgcgg gctgaccctg 60  
 cgggcccgcct ggccacctcg gccggctctc cacgaatcct ccttggcctg ccggcggggcg 120  
 ttacagtttc gggaagtgat tttgctcgtg ttccgaatgc aggctagcgg tgttcctggg 180  
 gcggttgggc aggtcccagc aacagtgggtg atatccctca taagggcgaa gcgacttcgt 240  
 cacgttgcgt aatgcgggat cctgcttcgt agctcgggtg gtcattgccag actgcgcacg 300  
 cggacctgca gcgggcccgcg aaatcccggc gaggaagggc gcgatgcgga ttctggtcac 360  
 cggcggagcc ggtttcatcg gctcgcaacta cgttcggcag ttgctcgggtg gtgcgtaccc 420  
 cgcattcggc gacgccgacg tggtcgtgct cgacaagctc acctacgccg gcaacgaggc 480  
 gaacctggcg ccggtcgcgg acaacccccg gctgaagttc gtctgcggcg acatctgcga 540

ccgcgaactg gttggcgggc tgatgtccgg cgtggacgtg gtggtgcaact tcgcgcgcga 600  
 aaccacagtc gaccgctcga tcaccggctc ggacgccttc gtgatcacca acgtgggtcgg 660  
 caccaacgtg ctgctgcagg ccgcgctcga cgccgagatc ggcaagtctg tgcacgtttc 720  
 caccgacgag gtctacggct ccatcgagga cggctcgtgg ccggaagacc acgcgctgga 780  
 gccgaattcc ccgtactcgg cggcgaaagc gggctcggac ctgctggccc gcgcctacca 840  
 ccgcacccac ggactgccgg tgtgcatcac ccgctgctcc aacaactacg ggccctacca 900  
 gttcccgag aaggtgctgc cgctgttcat cacgaacctg atggacggca gccaggtgcc 960  
 gctctacggc gacgggtca acgtgcggga ctggctgcac gtcagcgacc actgccgggg 1020  
 catccagctg gtggccgact ccgggcgcgc gggcgagatc tacaacatcg gcggcgccac 1080  
 cgagctgacc aacaacgagc tgaccgagcg gctgctggca gagctgggccc tcgactggtc 1140  
 ggtggtgctg ccggtcaccg accgcaaggg ccacgaccgc cgctactcgg tggaccacag 1200  
 caagatcgtc gaggaactgg ggtacgcgcc gcaggtcgac ttcgagaccg ggctgcgcga 1260  
 gacaatccgc tggtagcagg acaaccggga ctggtgggag ccgctgaagg cccgatcggc 1320  
 ggtggctcga tgagtgcct cgccgtgctg gtgcccggcg gccgcggcca gctgggctcg 1380  
 gagctggccc ggatcctccc cgcgcgagc ggggcgctgg tgcaccggcc gggttccggg 1440  
 gaactggacg tcaccgacgc cgaggaggtc gccgacgcgt tgggttcctt cgcggagacg 1500  
 gcgaaggacg cggagctgct accggtggtg atcaacgccg cggcgtagac gccggtggac 1560  
 gcggccgagt ccgaccggga ccgcgcggcc cggatcaacg ccgaaggcgc ggcctcgctg 1620  
 gcgaaagcgt gccggagcag cggctctgcc ctggtgcacg tgcgacgga ttacgtgttc 1680  
 cccggtgatg gggcccggcc gtacgagccg acggaccgga ccgggcgcg atcggtctac 1740  
 gggcgcacca agctcgaagg cgaacgggcc gtgctggagt ccggcgcgcg ggcctgggtg 1800  
 gtgcgcacgg catgggtgta cggcgcgagc ggcaagaact tcctgaaaac gatgatccgc 1860  
 ctctcggggg agcgcgacac gctgtccgtt gtggacgac agatcggtc gccgacttgg 1920  
 gcggcgagacc tggcgagcgg cctgctggag ctggccgaac gggtcgcga acgccgtgga 1980  
 ccggagcaga aggtgctgca ctgcaccaat tccggccagg tgacctggtg cgagtctcgc 2040  
 cgggcgatct tcgcggaatt cggcctggac gagaaccgcg tccaccctg cacgacggcg 2100  
 gacttcccc tcccggcga ccgcccggcc tactcggtc tgtccgacgt ggcgtggcga 2160  
 gaggcggggc tgaccccgat gcgcacctg cgggaagccc tggcgggcgc cttcgagaaa 2220  
 gacggcgaac ccctccgaac ccgctgacca gtcaccggga gggcgcgagt agccccggca 2280  
 gggccgcttc gacgcgatat cggctggcgc ggtgcgcaca atgggtgtcg ccggggcgag 2340  
 gaaggaaggc caggtgcccc ggg 2363

<210> 52

<211> 990

<212> DNA

<213> *Saccharopolyspora spinosa*

<220>

<221> CDS

<222> (1)..(987)

<223> ORF23; dNTP-Glucose-4,6-Dehydratase

<400> 52

atg cgg att ctg gtc acc ggc gga gcc ggt ttc atc ggc tgc cac tac 48  
 Met Arg Ile Leu Val Thr Gly Gly Ala Gly Phe Ile Gly Ser His Tyr  
 1 5 10 15

gtt cgg cag ttg ctc ggt ggt gcg tac ccc gca ttc gcc gac gcc gac 96  
 Val Arg Gln Leu Leu Gly Gly Ala Tyr Pro Ala Phe Ala Asp Ala Asp  
 20 25 30

gtg gtc gtg ctc gac aag ctc acc tac gcc ggc aac gag gcg aac ctg 144  
 Val Val Val Leu Asp Lys Leu Thr Tyr Ala Gly Asn Glu Ala Asn Leu  
 35 40 45

gcg ccg gtc gcg gac aac ccc cgg ctg aag ttc gtc tgc ggc gac atc 192  
 Ala Pro Val Ala Asp Asn Pro Arg Leu Lys Phe Val Cys Gly Asp Ile  
 50 55 60

tgc gac cgc gaa ctg gtt ggc ggc ctg atg tcc ggc gtg gac gtg gtg 240  
 Cys Asp Arg Glu Leu Val Gly Gly Leu Met Ser Gly Val Asp Val Val  
 65 70 75 80

gtg cac ttc gcc gcc gaa acc cac gtc gac cgc tcg atc acc ggc tcg 288  
 Val His Phe Ala Ala Glu Thr His Val Asp Arg Ser Ile Thr Gly Ser  
 85 90 95

gac gcc ttc gtg atc acc aac gtg gtc ggc acc aac gtg ctg ctg cag 336  
 Asp Ala Phe Val Ile Thr Asn Val Val Gly Thr Asn Val Leu Leu Gln  
 100 105 110

gcc gcg ctc gac gcc gag atc ggc aag ttc gtg cac gtt tcc acc gac 384  
 Ala Ala Leu Asp Ala Glu Ile Gly Lys Phe Val His Val Ser Thr Asp  
 115 120 125

gag gtc tac ggc tcc atc gag gac ggc tcg tgg ccc gaa gac cac gcg 432  
 Glu Val Tyr Gly Ser Ile Glu Asp Gly Ser Trp Pro Glu Asp His Ala  
 130 135 140

ctg gag ccg aat tcc ccg tac tcg gcg gcg aaa gcg ggc tcg gac ctg 480  
 Leu Glu Pro Asn Ser Pro Tyr Ser Ala Ala Lys Ala Gly Ser Asp Leu  
 145 150 155 160

ctg gcc cgc gcc tac cac cgc acc cac gga ctg ccg gtg tgc atc acc 528  
 Leu Ala Arg Ala Tyr His Arg Thr His Gly Leu Pro Val Cys Ile Thr  
 165 170 175

cgc tgc tcc aac aac tac ggg ccc tac cag ttc ccg gag aag gtg ctg 576  
 Arg Cys Ser Asn Asn Tyr Gly Pro Tyr Gln Phe Pro Glu Lys Val Leu  
 180 185 190

ccg ctg ttc atc acg aac ctg atg gac ggc agc cag gtg ccg ctc tac 624  
 Pro Leu Phe Ile Thr Asn Leu Met Asp Gly Ser Gln Val Pro Leu Tyr  
 195 200 205

ggc gac ggg ctc aac gtg cgg gac tgg ctg cac gtc agc gac cac tgc 672  
 Gly Asp Gly Leu Asn Val Arg Asp Trp Leu His Val Ser Asp His Cys  
 210 215 220

cgg ggc atc cag ctg gtg gcc gac tcc ggg cgc gcg ggc gag atc tac 720  
 Arg Gly Ile Gln Leu Val Ala Asp Ser Gly Arg Ala Gly Glu Ile Tyr  
 225 230 235 240

aac atc ggc ggc ggc acc gag ctg acc aac aac gag ctg acc gag cgg 768  
 Asn Ile Gly Gly Gly Thr Glu Leu Thr Asn Asn Glu Leu Thr Glu Arg  
 245 250 255

ctg ctg gca gag ctg ggc ctc gac tgg tcc gtg gtg cgg ccg gtc acc 816  
 Leu Leu Ala Glu Leu Gly Leu Asp Trp Ser Val Val Arg Pro Val Thr  
 260 265 270

gac cgc aag ggc cac gac cgc cgc tac tcc gtg gac cac agc aag atc 864  
 Asp Arg Lys Gly His Asp Arg Arg Tyr Ser Val Asp His Ser Lys Ile  
 275 280 285

gtc gag gaa ctg ggg tac gcg ccg cag gtc gac ttc gag acc ggg ctg 912  
 Val Glu Glu Leu Gly Tyr Ala Pro Gln Val Asp Phe Glu Thr Gly Leu  
 290 295 300

cgc gag aca atc cgc tgg tac cag gac aac cgg gac tgg tgg gag ccg 960  
 Arg Glu Thr Ile Arg Trp Tyr Gln Asp Asn Arg Asp Trp Trp Glu Pro  
 305 310 315 320

ctg aag gcc cga tcc gcg gtg gct cga tga 990  
 Leu Lys Ala Arg Ser Ala Val Ala Arg  
 325

<210> 53  
 <211> 329  
 <212> PRT  
 <213> Saccharopolyspora spinosa

<400> 53  
 Met Arg Ile Leu Val Thr Gly Gly Ala Gly Phe Ile Gly Ser His Tyr  
 1 5 10 15

Val Arg Gln Leu Leu Gly Gly Ala Tyr Pro Ala Phe Ala Asp Ala Asp  
 20 25 30

Val Val Val Leu Asp Lys Leu Thr Tyr Ala Gly Asn Glu Ala Asn Leu  
 35 40 45



Ala Pro Val Ala Asp Asn Pro Arg Leu Lys Phe Val Cys Gly Asp Ile  
 50 55 60

Cys Asp Arg Glu Leu Val Gly Gly Leu Met Ser Gly Val Asp Val Val  
 65 70 75 80

Val His Phe Ala Ala Glu Thr His Val Asp Arg Ser Ile Thr Gly Ser  
 85 90 95

Asp Ala Phe Val Ile Thr Asn Val Val Gly Thr Asn Val Leu Leu Gln  
 100 105 110

Ala Ala Leu Asp Ala Glu Ile Gly Lys Phe Val His Val Ser Thr Asp  
 115 120 125

Glu Val Tyr Gly Ser Ile Glu Asp Gly Ser Trp Pro Glu Asp His Ala  
 130 135 140

Leu Glu Pro Asn Ser Pro Tyr Ser Ala Ala Lys Ala Gly Ser Asp Leu  
 145 150 155 160

Leu Ala Arg Ala Tyr His Arg Thr His Gly Leu Pro Val Cys Ile Thr  
 165 170 175

Arg Cys Ser Asn Asn Tyr Gly Pro Tyr Gln Phe Pro Glu Lys Val Leu  
 180 185 190

Pro Leu Phe Ile Thr Asn Leu Met Asp Gly Ser Gln Val Pro Leu Tyr  
 195 200 205

Gly Asp Gly Leu Asn Val Arg Asp Trp Leu His Val Ser Asp His Cys  
 210 215 220

Arg Gly Ile Gln Leu Val Ala Asp Ser Gly Arg Ala Gly Glu Ile Tyr  
 225 230 235 240

Asn Ile Gly Gly Gly Thr Glu Leu Thr Asn Asn Glu Leu Thr Glu Arg  
 245 250 255

Leu Leu Ala Glu Leu Gly Leu Asp Trp Ser Val Val Arg Pro Val Thr  
 260 265 270

Asp Arg Lys Gly His Asp Arg Arg Tyr Ser Val Asp His Ser Lys Ile  
 275 280 285

Val Glu Glu Leu Gly Tyr Ala Pro Gln Val Asp Phe Glu Thr Gly Leu  
 290 295 300

Arg Glu Thr Ile Arg Trp Tyr Gln Asp Asn Arg Asp Trp Trp Glu Pro  
 305 310 315 320

Leu Lys Ala Arg Ser Ala Val Ala Arg  
 325

<210> 54

<211> 918

<212> DNA

<213> Saccharopolyspora spinosa

<220>

<221> CDS

<222> (1)..(915)

<223> ORF24; dNTP-4-keto-6-Deoxyglucose-3,5-Epimerase

<400> 54

atg agt cgc ctc gcc gtg ctg gtg ccc ggc ggc cgc ggc cag ctg ggc 48  
 Met Ser Arg Leu Ala Val Leu Val Pro Gly Gly Arg Gly Gln Leu Gly  
 1 5 10 15

tcg gag ctg gcc cgg atc ctc ccc gcg cgg acg ggg gcg ctg gtg cac 96  
 Ser Glu Leu Ala Arg Ile Leu Pro Ala Arg Thr Gly Ala Leu Val His  
 20 25 30

cgg ccg ggt tcc ggg gaa ctg gac gtc acc gac gcc gag gag gtc gcc 144  
 Arg Pro Gly Ser Gly Glu Leu Asp Val Thr Asp Ala Glu Glu Val Ala  
 35 40 45

gac gcg ttg ggt tcc ttc gcg gag acg gcg aag gac gcg gag ctg cga 192  
 Asp Ala Leu Gly Ser Phe Ala Glu Thr Ala Lys Asp Ala Glu Leu Arg  
 50 55 60

ccg gtg gtg atc aac gcc gcg gcg tac acg gcg gtg gac gcg gcc gag 240  
 Pro Val Val Ile Asn Ala Ala Tyr Thr Ala Val Asp Ala Ala Glu  
 65 70 75 80

tcc gac ccg gac cgc gcg gcc cgg atc aac gcc gaa ggc gcg gcc tcg 288  
 Ser Asp Pro Asp Arg Ala Ala Arg Ile Asn Ala Glu Gly Ala Ala Ser  
 85 90 95

ctg gcg aaa gcg tgc cgg agc agc ggt ctg ccc ctg gtg cac gtg tcg 336  
 Leu Ala Lys Ala Cys Arg Ser Ser Gly Leu Pro Leu Val His Val Ser  
 100 105 110

acg gat tac gtg ttc ccc ggt gat ggg gcc cgg ccg tac gag ccg acg 384

Thr Asp Tyr Val Phe Pro Gly Asp Gly Ala Arg Pro Tyr Glu Pro Thr	
115 120 125	
gac ccg acc ggg ccg cga tcg gtc tac ggg cgc acc aag ctc gaa ggc	432
Asp Pro Thr Gly Pro Arg Ser Val Tyr Gly Arg Thr Lys Leu Glu Gly	
130 135 140	
gaa cgg gcc gtg ctg gag tcc ggc gcg cgg gcc tgg gtg gtg cgc acg	480
Glu Arg Ala Val Leu Glu Ser Gly Ala Arg Ala Trp Val Val Arg Thr	
145 150 155 160	
gca tgg gtg tac ggc gcg agc ggc aag aac ttc ctg aaa acg atg atc	528
Ala Trp Val Tyr Gly Ala Ser Gly Lys Asn Phe Leu Lys Thr Met Ile	
165 170 175	
cgc ctc tcg ggg gag cgc gac acg ctg tcc gtt gtg gac gat cag atc	576
Arg Leu Ser Gly Glu Arg Asp Thr Leu Ser Val Val Asp Asp Gln Ile	
180 185 190	
ggc tcg ccg act tgg gcg gcg gac ctg gcg agc ggc ctg ctg gag ctg	624
Gly Ser Pro Thr Trp Ala Ala Asp Leu Ala Ser Gly Leu Leu Glu Leu	
195 200 205	
gcc gaa cgg gtc gcc gaa cgc cgt gga ccg gag cag aag gtg ctg cac	672
Ala Glu Arg Val Ala Glu Arg Arg Gly Pro Glu Gln Lys Val Leu His	
210 215 220	
tgc acc aat tcc ggc cag gtg acc tgg tac gag ttc gcg cgg gcg atc	720
Cys Thr Asn Ser Gly Gln Val Thr Trp Tyr Glu Phe Ala Arg Ala Ile	
225 230 235 240	
ttc gcg gaa ttc ggc ctg gac gag aac cgc gtc cac ccg tgc acg acg	768
Phe Ala Glu Phe Gly Leu Asp Glu Asn Arg Val His Pro Cys Thr Thr	
245 250 255	
gcg gac ttc ccc ctc ccg gcg cac cgc ccg gcc tac tcg gtc ctg tcc	816
Ala Asp Phe Pro Leu Pro Ala His Arg Pro Ala Tyr Ser Val Leu Ser	
260 265 270	
gac gtg gcg tgg cga gag gcg ggc ctg acc ccg atg cgc acc tgg cgg	864
Asp Val Ala Trp Arg Glu Ala Gly Leu Thr Pro Met Arg Thr Trp Arg	
275 280 285	
gaa gcc ctg gcg gcg gcc ttc gag aaa gac ggc gaa acc ctc cga acc	912
Glu Ala Leu Ala Ala Phe Glu Lys Asp Gly Glu Thr Leu Arg Thr	
290 295 300	
cgc tga	918

Arg  
305

<210> 55  
<211> 305  
<212> PRT  
<213> Saccharopolyspora spinosa

<400> 55

Met Ser Arg Leu Ala Val Leu Val Pro Gly Gly Arg Gly Gln Leu Gly  
1 5 10 15

Ser Glu Leu Ala Arg Ile Leu Pro Ala Arg Thr Gly Ala Leu Val His  
20 25 30

Arg Pro Gly Ser Gly Glu Leu Asp Val Thr Asp Ala Glu Glu Val Ala  
35 40 45

Asp Ala Leu Gly Ser Phe Ala Glu Thr Ala Lys Asp Ala Glu Leu Arg  
50 55 60

Pro Val Val Ile Asn Ala Ala Ala Tyr Thr Ala Val Asp Ala Ala Glu  
65 70 75 80

Ser Asp Pro Asp Arg Ala Ala Arg Ile Asn Ala Glu Gly Ala Ala Ser  
85 90 95

Leu Ala Lys Ala Cys Arg Ser Ser Gly Leu Pro Leu Val His Val Ser  
100 105 110

Thr Asp Tyr Val Phe Pro Gly Asp Gly Ala Arg Pro Tyr Glu Pro Thr  
115 120 125

Asp Pro Thr Gly Pro Arg Ser Val Tyr Gly Arg Thr Lys Leu Glu Gly  
130 135 140

Glu Arg Ala Val Leu Glu Ser Gly Ala Arg Ala Trp Val Val Arg Thr  
145 150 155 160

Ala Trp Val Tyr Gly Ala Ser Gly Lys Asn Phe Leu Lys Thr Met Ile  
165 170 175

Arg Leu Ser Gly Glu Arg Asp Thr Leu Ser Val Val Asp Asp Gln Ile  
180 185 190

Gly Ser Pro Thr Trp Ala Ala Asp Leu Ala Ser Gly Leu Leu Glu Leu  
195 200 205

Ala Glu Arg Val Ala Glu Arg Arg Gly Pro Glu Gln Lys Val Leu His  
210 215 220

Cys Thr Asn Ser Gly Gln Val Thr Trp Tyr Glu Phe Ala Arg Ala Ile  
225 230 235 240

Phe Ala Glu Phe Gly Leu Asp Glu Asn Arg Val His Pro Cys Thr Thr  
245 250 255

Ala Asp Phe Pro Leu Pro Ala His Arg Pro Ala Tyr Ser Val Leu Ser  
260 265 270

Asp Val Ala Trp Arg Glu Ala Gly Leu Thr Pro Met Arg Thr Trp Arg  
275 280 285

Glu Ala Leu Ala Ala Ala Phe Glu Lys Asp Gly Glu Thr Leu Arg Thr  
290 295 300

Arg  
305

**THIS PAGE BLANK (USPTO)**